A proposal for a
Chechen orthography
Erwin R. Komen

Abstract

Chechen is an intriguing Caucasian language having a rich consonantal inventory as well as an extremely large vowel inventory.

In this paper the phonemic system of the language is reviewed and compared with the currently used Cyrillic orthography. The main drawback of this orthography is, that it does not provide enough symbols to discern all vowel phonemes.

In light of the fact that an invitation was issued to come up with an orthography proposal, two proposals are given in this paper. If it is obligatory that Chechen be written with an orthography based on the Cyrillic alphabet, then the first proposal describes the measures that could be undertaken to reform the orthography. The second proposal describes an orthography based on the Latin script. When purely academic considerations are taken into account, it is argued that for the Chechen language a Latin based orthography has advantages over a Cyrillic based one. A form of the described Latin based orthography has for more than ten years been used by scholars.

1. Introduction

Chechen is a language spoken in the Caucasus by approximately one million people. The sound system of the language is characterized by a large number of different vowel phonemes, and as for the consonants – like many languages in the general Caucasus area it has ejectives, as well as an epiglottal stop and fricative (comparable to the Arabic "ayin").

The Chechen language has a relatively short but complex history of orthography. In the 19th century the Arabic script was used in personal communication between the Chechens living in their North-Caucasian homeland and those that had moved house to the Ottoman empire (nowadays Turkey, Jordan and Syria). At the end of the 19th century the renown baron Uslar, who studied many Caucasian languages, devised an orthography using Cyrillic as well as self-invented characters. This orthography was used only in his scientific work. Later on, in the Lenin era of the Soviet Union, in the 1920's, scholars created a Latin-based orthography. Primers, books and newspapers used this orthography. Then, as the policy changed under Stalin in the 1930's, this orthography was changed into a Cyrillic one. From that time on the Cyrillic orthography has been used as the main orthography.

Literacy rates in the mother tongue have not been high. This is probably due to a number of factors. One of these factors has been the quality of the orthography itself. In this paper I would like to review the problems met with the current orthography, and come up with two new proposals for a new orthography. One proposal will be to use a Latin-based orthography, which I believe that would be the best in the long run for this language and for its speakers, wherever they are. The other proposal will be a Cyrillic-based one, since other factors might not allow switching to a Latin-based orthography.

This paper is organized as follows. In section 2 I will review the sound system of the Chechen language, focusing on the difference between phones and phonemes, and also focusing on phonemic contrasts. Section 3 will review the limitations of the current Cyrillic based orthography, which to a large part can be understood against the background of the sound system described in section 2. Proposals for linguistically better orthographies are given in section 5). The proposals are examplified using a small story in section 5. Finally section 6 summarizes the conclusions and contains a discussion as to how to proceed.

2. The sound system of Chechen

In this section I would like to explain the *system* of the Chechen sounds. Such a system is based on phonemes rather than phones, on phonemic oppositions rather than phonetic ones. The explanation here is based on previous research (Nichols 1994, 1997, Komen 1993, 1996, 2007).

2.1. Phonemic opposition

The consonant and the vowel inventory of a language consists of phonemes: sounds that are in contrastive opposition to one another. In order to understand the concept of the phoneme, look at the two words given in (1) and (2). The sounds [tʼ] and [d] are in contrastive opposition, because they are part of a minimal pair[1]. The environments

---

[1] A minimal pair is a set of two words that minimally differ in one sound, yet have a different meaning. The different sounds should be phonetically "close" to one another. The sounds [t] and [d] for instance are close to one another, since they only differ in voicing.

of these two consonants are exactly the same: both occur at the beginning of a word, and both are followed by the long vowel written here as 'aː'.

(1) [tʼaː]　　'forefoot'

(2) [daː]　　'dad'

When two phones, or distinct sounds, are in contrastive opposition, then they are considered to belong to distinct phonemes. A phoneme is identified by writing it between slashes, as in /tʼ/ and /d/.

One phoneme can have different realizations (phones), depending on the conditioning environment, the environment in which they occur, and which somehow help to determine how the phoneme is pronounced exactly.

This can be illustrated by looking at the phoneme /r/. Normally the trill is realized by the phone [r], as in (3) where it occurs word-finally.

(3) [kir]　　'whitewash'

However, when the phoneme /r/ occurs between vowels (which is usually called "intervocalically"), then the native speakers I consulted realize this phoneme as a *flap*, i.e. as a trill that lasts only one beat, as for example in (4).

(4) [hɜɾɜ]　　'this'

So we can conclude that the phoneme /r/ has two realizations: intervocalically it is realized as a flap [ɾ], while in all other cases it occurs as a trill [r].

One important conclusion should be drawn from the difference between phones and phonemes above:

(5) **Phoneme conclusion**
Since the different phones of a phoneme occur in different environments, and therefore never contrast, they can be written in the orthography with the same grapheme (i.e. a letter or digraph or so).

For the reason stated above it is of primary importance to identify what exactly the phonemes are of a language. Every phoneme may be realized by different phones, and these phones should never overlap – that is to say: no phone can be part of two different phonemes. So, since the amount of phonemes is almost always smaller than the amount of phones, a language can contain a large amount of sounds (phones), but still only need a limited amount of graphemes to represent the phonemes faithfully.

2.2. The consonants

In this section I will look at the consonant inventory of the Chechen language. I will do this on the basis of earlier research, and I will *not* offer a new analysis based on current research.

There may not be a general consensus as to the number of consonant phonemes the Chechen language has, but, as I will show, the differences in opinion do not necessarily lead to differences in orthography.

One of the first phonemic inventories of the language is given in Table 1 (Komen 1993). According to this inventory a distinction is made between consonant phonemes in three different ways. Every row in the table constitutes a different *place of articulation* for the phoneme. Note that a phonemic place of articulation does not necessarily coincide completely with a phonetic place of articulation. For instance the

phoneme /t/ is identified as "alveolar", but its realizations could all actually have a "dental" place of articulation. In a phonemic inventory as shown here the places of articulation are taken together where possible and where allowable so as to get more general categories, abstracting away a little bit more from reality.

The columns in this inventory are grouped together around larger areas, identifying the *manner of articulation*. Plosives are sounds having a complete stricture in the mouth area. Fricatives have a very tiny stricture. Affricates start out with a stricture, but give way to a tiny stricture very soon. Laterals have a stricture in the middle, but not at the sides. Trills have a stricture periodically. And glides have a large stricture. Usually they are phonetically the same as vowels. Only within the whole system of the language can a decision be made when and whether glides are to be regarded as vowels or consonants.

The groups of columns have been subdivided into smaller columns, depending on additional qualities of the sound. The phones can be voiced or voiceless. Plosives and affricates can also be ejectivized. Some consonants are phonemically long (also labeled as "fortis"). Furthermore there are nasals.

Table 1 Consonant phonemes according to Komen (1993)

	Plosive					Fricative				Affricate		Lateral	Trill		Glide
	Ejec	Vl	Long	Vd	Ns	Vl	Long	Vd	Long	Ejec	Vl		Vd	Vl	
Labial	/p'/	/p/	/p:/	/b/	/m/										/w/
Alveolar	/t'/	/t/	/t:/	/d/	/n/	/s/	/s:/	/z/	/z:/	/ts'/	/ts/	/l/	/r/	/r̥/	
Post Alveolar						/ʃ/		/ʒ/		/tʃ'/	/tʃ/				/j/
Velar	/k'/	/k/	/k:/	/g/											
Uvular	/q'/	/q/	/q:/			/χ/		/ʁ/							
Pharyngeal						/ħ/									
Glottal		/ʔ/		/ʕ/	/h/										

An other subdivision of the Chechen phonemes is shown in Table 2 (Nichols 1994, 1997). In this division no affricates are recognized as a separate group. Instead a further distinction is made in the place of articulation. While the first research made a distinction only between "Alveolar" and "Post-Alveolar", the subdivision according to Nichols shows a three-way distinction between "Dental", "Alveolar" and "Palatal".

The actual amount of phonemes identified is almost the same. Nichols' consonant phoneme inventory does have two more phonemes than the first inventory shown here. These are the voiced alveolar affricate /dz/ and the voiced palatal affricate /dʒ/. Nichols is very right in noting these phonemes. They are there in Chechen. However, they occur on a very limited scale, and their occurrence depends on the dialect spoken. The dialect underlying the inventory in Table 1 did not include a phonemic distinction between [z] and [dz], nor between [ʒ] and [dʒ].

Table 2 Consonant phonemes according to Nichols (1994, 1997)

	Noncontinuant Obstruents				Continuant Obstruents		Nasal	Resonants Liquid		Glide
	Vl	Vd	Ejec	Fortis	Vl	Vd		Vd	Vl	
Labial	/p/	/b/	(/p'/)	/pː/			/m/			/w/
Dental	/t/	/d/	/t'/	/tː/				/l/		
Alveolar	/ts/	/dz/	/ts'/	/sː/	/s/	/z/	/n/	/r/	/r̥/	
Palatal	/tʃ/	/dʒ/	/tʃ'/		/ʃ/	/ʒ/				/j/
Velar	/k/	/g/	/k'/	/xk/[2]	/x/					
Uvular	/q/		/q'/	/qː/		/ʁ/				
Pharyngeal		/ʕ/			/ħ/					
Glottal		/ʔ/			/h/					

As I will show further in section 3, the current orthography is perfectly capable of distinguishing all the consonant phonemes identified either by Komen or by Nichols. Unfortunately it doesn't do so in a consistent way. The proposals made in section 5) will show ways to improve consistency.

For the rest of this paper I will be using Nichols' way to show the phoneme inventory of consonants, since that shows more phonemes, and therefore potentially poses a larger challenge for an orthography.

2.3. The vowels

Vowels have traditionally been the most underrepresented phonemes in Chechen orthographies (Komen 1996). With underrepresentation I mean that one grapheme is used to represent distinctly different phonemes. The grapheme *o*, for example, can stand for the short /o/ phoneme, the long /oː/ phoneme, the short /uo/ or the long /uoː/. Underrepresentation of vowel phonemes may be one of the main causes why mother tongue speakers find it difficult to learn to read and write their own language.

The information given in this section is based on research conducted in one particular variant of the Chechen lowland's dialect (Komen 2007). The vowel phoneme inventory taken from that research is repeated here in Table 3.

---

[2] According to Nichols this is the "structural and etymological equivalent of *kk*.

Table 3 Vowel phoneme inventory

		Front				Back				Diphthong	
		Unr		Rnd		Unr		Rnd			
		Short	Long	Short	Long	Short	Long	Short	Long		
Close	plain	i	iː	y	yː			u	uː	ui	
	diphthong	ie	ieː	ye	yeː			uo	uoː	uoi	ou
Mid		e	eː					o	oː	ey	
Open						a	aː			ai	

Phonemically Chechen distinguishes between short and long vowels. Phonetically one will find at least three different "a" sounds in Chechen, but this doesn't mean that there are three phonemes needed. The short /a/ is quite short, it is realized as [ɜ], which is very close to the [ə]. The short /a/ only has one realization that is independent of the environment. However, the long /aː/ phoneme has at least two different realizations, depending on the conditioning environment. In closed syllables the long /aː/ is realized as an [a] phone. This phone is a bit longer than the short [ɜ], and its quality is different (the mouth is more open). When the long /aː/ occurs in an open syllable, it is realized as an [aː], which is much longer than the realization in a closed syllable.

The same story goes for the short /i/ and /u/ and their long counterparts /iː/ and /uː/. The three other short vowel phonemes /e/, /o/ and /y/ also have long counterparts /eː/, /oː/ and /yː/. However, these long counterparts do not occur word-finally.

Chechen also distinguishes between short and long diphthongs that start with a close vowel (i.e. the vowels /i/, /u/ or /y/). But the distinction between short and long is neutralized in closed syllables. On top of that, the short diphthongs only occur in open syllables that are not word-final. Sequences ending on a close vowel have all been interpreted as vowel diphthongs.

2.4. The sound [æ]

There is one particular sound, the [æ], which traditionally has had a grapheme in all orthographies encountered. Purely theoretically, this would not have been necessary. The normal [æ] is a realization of the long /eː/ phoneme in closed syllables. And the pharyngealized [æ̰] is a realization of the short /e/ phoneme where the conditioning environment contains a pharyngeal sound.

This is *not* to say that it is *bad* to have some overdifferentiation.

With overdifferentiation I mean that different graphemes (in this case cyrillic аь and е) are used to represent the same phoneme (in this case the long /eː/ phoneme).

3.   Current orthography

The current Cyrillic orthography allows all consonant phonemes to be expressed using separate graphemes (Dzhamalxanov 1972, 1992, Maciev 1961). However, there are some inconsistencies that make it difficult for beginning readers. The

vowels underrepresent the existing phonemes in an unsystematic way[3], leading to difficulties in reading, and even to ambiguities[4].

3.1. The consonants in the current orthography

Most of the consonants are expressed in a good and logical way using the current orthography, as can be seen from Table 4. Some notes are in place for the representation of individual consonant phonemes.

- The glottal stop is *not* represented by a grapheme when there only is a single glottal stop between vowels. So a word like /daːʔim/ 'always' is written as даим. But a *doubled* glottal stop (consonant geminating can be part of a morpheme, like for instance intensification) between vowels is represented by a hard-sign ъ, as for instance with a word like /daʔːalts/ 'until eating' даъалц. This hard-sign is also used to represent a *single* glottal stop at the end of a word, as in /tsħaʔ/ 'one' цхьаъ.
- The glide /j/ is represented by Cyrillic й when it comes syllable-final. But when it comes syllable-initial, then the representation depends on the following vowel, as for example: /ji/ = йи, /je/ = е, /ja/ = я, /jo/ = йо, /ju/ = ю, /jy/ = юь.
- Besides the *fortis* consonants other consonants can also be doubled. This is indicated by writing the first part of the grapheme twice. So a doubled /tː/ is written as ттӀ.

---

[3] From the 6 long vowels only the /iː/ and /yː/ have separate graphemes. From the 6 diphthongs only the /ye/ has a separate grapheme.

[4] Different kinds of ambiguities have been reported (Komen 1996). The most severe ones are those where the difference between the plain and iterative aspects of a set of verbs can not be resolved in the existing orthography. It is for instance not possible to distinguish between /letʼqʼa/ 'mourn.PN.PRS ', /lietʼqʼa/ 'mourn.IT.PRS'. Both are written as къеда.

Table 4 Consonant phonemes in current Cyrillic orthography

	Noncontinuant Obstruents				Continuant Obstruents		Nasal	Liquid		Glide
	VI	Vd	Ejec	Fortis	VI	Vd	Vd	Vd	VI	
Labial	/p/ п	/b/ б	(/p'/) пI	/pː/ пп			/m/ м			/w/ в
Dental	/t/ т	/d/ д	/t'/ тI	/tː/ тт				/l/ л		
Alveolar	/ts/ ц	/dz/ дз	/ts'/ цI	/sː/ сс	/s/ с	/z/ з	/n/ н	/r/ р	/r̥/ рхI	
Palatal										й я е
	/tʃ/ ч	/dʒ/ дж	/tʃ'/ чI		/ʃ/ ш	/ʒ/ ж				/j/ ю юь
Velar	/k/ к	/g/ г	/k'/ кI	/xk/ хк	/x/ х					
Uvular	/q/ кх		/q'/ къ	/qː/ ккх		/ʁ/ гI				
Pharyngeal		/ʕ/ I			/ħ/ хь					
Glottal		/ʔ/ -, ъ			/h/ хI					

3.2. Vowels in the current orthography

The representation of the vowel phonemes in the current Cyrillic orthography needs some explanation. A preliminary overview is given in Table 5. As to the writing system the following things should be noted:

- The Cyrillic vowel э is used to represent /e/, /eː/, /ie/ and /ieː/ when these occur syllable initial, as for instance in /sħaʔietsa/ 'to take' схьаэца. The Cyrillic vowel е is used to represent these same vowel phonemes in the following instances: (a) whenever preceded by a /j/, and (b) after consonants.
- Rarely in current writing, but generally in one dictionary (Maciev 1960) the long /eː/ or /ieː/ are, only in open syllables, written with an overbar.
- In the same way a distinction between long /aː/, /oː/ and short /a/, /o/ can be made with an overbar. But this is almost exclusively done in the 1960 dictionary.
- The diphthongs /ie/, /ieː/, /uo/ and /uoː/ are only very rarely written "full" as иэ and уо. Besides, this is only done in open syllables.

Table 5 Vowel phonemes in current Cyrillic orthography

	Front				Back				Diphthong	
	Unr		Rnd		Unr		Rnd			
	Short	Long	Short	Long	Short	Long	Short	Long		
Close	/i/ и	/i:/ ий	/y/ уь	/y:/ уьй			/u/ у	/u:/ у	/ui/ уй	
	/ie/ э е иэ	/ie:/ э е иэ	/ye/ оь	/ye:/ оь			/uo/ о уо	/uo:/ о уо	/uoi/ ой	/ou/ ов
Mid	/e/ э е аь	/e:/ э е аь					/o/ о	/o:/ о		/ey/ эв ев
Open					/a/ а	/a:/ а			/ai/ ай	

3.3. Other features of the current orthography

As explained by Seifart, orthographies do not only consist of a set of graphic symbols, but usually are also accompanied by rules (Seifart 2006:277). The current Cyrillic orthography has one particular orthographic rule that needs to be highlighted at this point. Besides the Cyrillic graphemes that are used to represent the language's phonemes, the current Cyrillic orthography also uses the hard-sign ъ and the soft sign ь as syllable dividers.

The usual syllable-divider is the hard-sign ъ. This is used in cases where the hard-sign would not be confused with another of its usages. An example where it is used is for instance /ʃeljella/ 'has become cold', which is represented as шелъелла. If no syllable divider were used, then an incorrect syllable division of the word would result.

Whenever there is possible confusion, the soft-sign is used instead, as for instance in /hallakxilla/ 'perished', which is written as хӀаллакьхилла. If the hard-sign were used, then an incorrect consonant къ would be read.

3.4. Problems in the current orthography

An experienced reader is able to learn to read and write Chechen with the current Cyrillic orthography, but for beginning readers there are several challenges. I would now like to focus on these problem areas.

3.4.1. K-sound confusion

As can be seen from section 3.1 the four "k-sounds" к, кх, кӀ, and къ (as well as the combination кьх that can occur in a word like хӀаллакьхилла, as illustrated in section 3.3) are easily confused by beginning readers. I think that the problem is twofold:

- While all other plosives require only one letter to be written, the uvular plosive /q/ is written with a digraph. This in itself is not too problematic.
- Five of the six ejectives are written as: consonant + Latin I. The exception is the uvular ejective /qˀ/, which is written as: consonant + hard-sign.

The combination of these two problems cause beginning readers to stumble and not perceive the written material well.

On the one hand good teaching methods could help beginning readers to discern the k-sounds better. On the other hand, if there were an orthography that could do

away with one or two of the problems that are identified, then that would make life easier for the readers.

3.4.2. Noun-class markers

One of the main auxiliaries, many Chechen verbs and several other Chechen words take a consonant class-prefix in order to signify agreement with an argument (usually a noun or a noun phrase). The consonants used for these class-prefixes are the following: /v/, /j/, /b/ and /d/. There are no problems for the class-prefixes /v/, /b/ and /d/. But the fourth class prefix, the /j/, is written differently depending on the following vowel. The basic differences are as follows:

/j/ + /i/ is written as йи.
/j/ + /e/ is written as е.
/j/ + /a/ is written as я.
/j/ + /o/ is written as йо.
/j/ + /u/ is written as ю.
/j/ + /y/ is written as юь.

From a grammatical point of view the class-markers are morphemes, and it would be better if they were identifiable individually. Since the current orthography combines the class marker and a vowel in four of the six basic situations sketched above, there is a problem.

Again this problem could be tackled with good teaching. Experienced readers will have automized the different ways in which a class-marker can surface. But this problem does make life harder for beginning readers. If an orthography improvement is undertaken, then it would be good (and not too difficult) to tackle this problem.

3.4.3. Vowel underdifferentiation

A point of great concern for the current orthography is the underdifferentiation of vowels. As shown in section 3.2 only the short front close vowels /i/ and /y/ (in the Cyrillic orthography и and уь) have a different representation for their long counterparts /iː/ and /yː/ (ий and уьй respectively). But the short vowels /e/, /a/, /o/ and /u/ (in the Cyrillic orthography е, а, о and у) don't have a different grapheme for the corresponding long vowels /eː/, /aː/, /oː/ and /uː/.

In practice this means that a word like дакъа can mean either 'corpse' or 'part', depending on whether the first vowel is understood to be short (yielding /daq'a/) or long (yielding /daːq'a/). Context is needed to disambiguate such words. This in turn makes reading harder – especially for the beginning reader.

As for the diphthongs – they are underdifferentiated too. From the six diphthongs starting with a high vowel (i.e: /ie/, /ieː/, /uo/, /uoː/, /ye/ and /yeː/) only two (/ye/ and /yeː/) are normally written out, using the grapheme оь. In normal writing the other diphthongs are indiscernible from the short or long vowels. So the Cyrillic vowels е and э could be understood as any of the following phonemes: /ie/, /ieː/, /e/, /eː/, as ilustrated in examples (1) to (4).

1) /dieza/ 'to love'   деза

2) /deza/ 'valuable'   деза

3) /diexa/ 'to request'   деха

4) /deexa/'long'   деха

Likewise the Cyrillic vowels о could be understood as any of the following phonemes: /o/, /oː/, /uo/, /uoː/, as illustrated in examples (5) until (8). Only in rare cases are the diphthongs spelled out as уо and иэ (but only in open syllables).

5) /uoza/ 'to pull'   оза

6) /ooza/ 'slim'   оза

7) /dotta/ 'bake.PRS'   дотта

8) /duotta/ 'to pour'   дотта

4.   Orthography proposals

Recently an invitation by professor Timaev of the Chechen State University to anyone interested was issued via the radio to come up with proposals for an orthography and a spelling reform (Timaev 2007).

In light of that invitation this section contains two proposals for an orthography that deal with the problems encountered in the current one. The first proposal for a linguistically better orthography is given in section 4.1. This first proposal is Cyrillic-based. Finally section 4.2 focuses on a linguistically better orthography proposal, which is based on the Latin script.

4.1. Cyrillic-based orthography proposal

Before presenting a proposal for a new (or renewed) Cyrillic orthography, some considerations will be listed in section 4.1.1. Next in sections 4.1.2 and 4.1.3 I discuss several options, which for various reasons I would not recommend. Finally in section 4.1.4 the new proposal is presented and illustrated.

4.1.1. Considerations

The first question I would like to ask here, is whether a new Cyrillic based orthography could choose to completely redefine the graphemes used for several consonants and vowels or whether there should be minimal changes, while nothing crucial is changed from the existing orthography.

    The answer to this question depends on how many people would be affected by the choices made. Unfortunately, figures of the mother-tongue literacy rate of Chechen people are not published, as far as I am aware. The estimate that is made by some native speakers I know is very low. In their opinion probably only 5% of the total population of Chechens is able to freely read and write in their mother tongue. This percentage includes scholars and teachers – people with a higher education. It also includes individuals who kept the written language up out of their own interest. Part of the generation of youngsters between 15-25 communicate with one another over the internet. They use whatever symbols and ways they can think of to express themselves in Chechen (including Latin-based orthographies). The youth within Chechnya learns read and write in the current Cyrillic-based orthography, using older textbooks or using blackboard. Large part of the youth of the refugee population does

not learn to read and write their mother tongue. Small part does learn it, being educated in small groups.

Since a relatively small part of the population would be negatively impacted by drastic changes, the benefits of making a scientifically sound but drastic changes should outweigh the negative effect of people having to switch to yet another orthography.

But of course, if changes *can* be kept minimal, that would always be the preference!

If for instance the writing system of the vowels is changed, then the question is whether there are examples of other Cyrillic-based orthographies containing many vowels. Well, the answer to that is: "Yes, but…". Related or neighboring language groups in the Caucasus do *not* contain as many vowels as Chechen does. The closely related Ingush language does have many vowels, but has not yet considered an orthography change. The further related Batsby language is spoken in Georgia by only a few thousand people, and their orthography does not offer much help. Other languages from the Nakh-Dagestanian family are not as vowel-rich as the Nakh branch (containing Chechen, Ingush and Batsby). There are also several languages from the Turkic family spoken in the Caucasus (e.g. Kumyk). But, as far as vowels are concerned, languages from this family use additional Cyrillic symbols to represent the different vowel qualities needed. They don't need to represent length or diphthongs that substantially.

As far as Cyrillic-based orthographies are concerned, only languages from the Finno-Ugric family make consistent phonemic difference between short and long vowels. In some of their orthographies long vowels are represented using two dots above the grapheme. Other orthographies use the overbar – but only in dictionaries, as in Chechen.

In light of the fact that a diacritic has been available to signify length in Chechen, but has generally not been used, it may not be such a good idea to use any other diacritic.

4.1.2. Options for the k-sounds

Within the Cyrillic-based orthography one way to make a more consistent spelling of the k-sounds would be to write the uvular ejective /q'/ as a trigraph. The first part of the trigraph would be the normal way in which the uvular plosive /q/ is written, i.e. as a кх. The second part of the trigraph would be the normal symbol used in all other cases to represent ejectivization. The result would be that the /q'/ is written as кхӀ. Note that the sequence хӀ normally represents the /h/. But in Chechen there do not seem to be words where the combination /q'h/ occurs, so that the representation кхӀ is indeed unambiguous.

A <u>second option</u> would be more radical. There is one consonant grapheme unused in Chechen, except for Russian loan words, and that is the ф. This grapheme could in principle be used to represent the uvular plosive /q/, so that the logical way to write the uvular ejective /q'/ would become фӀ.

However, this would lead to a situation where Russian loan words are misread. And besides, even though this letter is not used in Chechen, it is used in the closely related Ingush language. This touches on the question how far the orthography of the Chechen and Ingush languages should be mutually intelligible.

On the whole, neither of the options offered above are very attractive. If Chechen has to be written with a Cyrillic-based orthography, then it might be better to leave the writing of the k-sounds as it is right now, and improve on the teaching methods. If Chechen is written using a Latin-based orthography, there won't be any problem with writing the k-sounds consistently (see section 4.2.1).

4.1.3. Options for the vowels

Since the usage of diacritics, as discussed above, has not proven itself to be as widely accepted as should be, there are few options left. I would like to discuss four options, and recommend one of them.

Note that all options discussed over here take as a starting point that the Cyrillic vowels я and ю are not used anymore in Chechen words (they might still be used in Russian loan words).

The first option would be the one where the least changes with respect to the current orthography are taken. While almost all the vowel graphemes could stay as they are right now, three accent signs could be used to signal that the vowel is short, that it is long, or that it is a diphthong. The accent signs would not have to be used all the time. On option would be to only use them in cases of potential ambiguity. Another option would be that a vowel is short, unless an accent sign indicates that it is long or that it is a diphthong. Early learners would need the accent signs more than advanced readers.

The six short vowel phonemes could be represented by а, е, и, о, у and уь, or their accented counterparts à, è, и, ò, ỳ and ỳь (note that there is no accented и available). Their long counterparts could then be represented using the following symbols: á, é, ий, ó, ý and ýь (or their unaccented counterparts if the length does not need to be indicated). The symbol оь would be retained, but a finer distinction between short and long would be possible using òь and óь. The remaining diphtongs would be indicated by ê for the short /ie/ and the long /ieː/, and the ô for the short /uo/ and the long /uoː/. For these diphthongs it would not be possible to distinguish the short ones from the long ones.

As attractive as this option may look like, there are several problems. The most obvious one is related to the usage of diacritic signs. The overbar diacritic of the existing orthography has not been used widely, so it seems unlikely that any diacritics will find broad acception. The second problem is that this option retains old inconsistencies. Five short vowel phonemes can be distinguished from their long counterparts using accent signs, whereas one short phoneme (the /i/ phoneme, written as и) denotes this differently (i.e. by adding a й symbol). It also introduces a new inconsistency: two short diphthongs are written using a diacritic and cannot be distinguished as to length, while one diphthong (the /ye/ phoneme, written as оь) is written without diacritic and can be distinguished in length.

Since – from the Cyrillic based orthographies – this option is closest to the existing orthography, the sample text in section 10.1 has been reproduced using this option in section 10.2.

The second option would be to use different symbols for the long and short vowels. The six short vowel phonemes could be represented by а, э, и, о, у and уь. Their long counterparts could then be represented using the other available Cyrillic vowel symbols: я, е, ы, ё, ю and юь.

Although this is possible, it does not seem to be very wise, since consistency is lacking. For four of the 6 graphemes the correspondence between short and long is more or less logical: а-я, э-е, у-ю and уь-юь. But the correspondency between the remaining two vowels doesn't really match: и-ы and о-ё.

<u>The third option</u> would be to use one sign *after* a vowel signifying vowel-length consistently for all 6 vowel phonemes. The most logical sign to be used would be the soft-sign, as exemplified in Table 6.

Table 6 Vowels with following length sign

		Front						Back			
		Unr		Rnd		Unr		Rnd			
		Short	Long	Short	Long	Short	Long	Short		Long	
Close	plain	/i/ и	/iː/ иь	/y/ ы	/yː/ ыь			/u/ у		/uː/ уь	
	diphthong	/ie/ иэ	/ieː/ иэ	/ye/ ые	/yeː/ ые			/uo/ уо		/uoː/ уо	
Mid		/e/ э	/eː/ эь					/o/ о		/oː/ оь	
Open						/a/ а	/aː/ аь				

This system radically differs from the existing vowel representations, leading to confusion for the few people who are now able to read and write. The soft-sign ь was until now used to differentiate between vowel quality (the symbol аь representing [æ], the symbol оь representing /ye/, and the symbol уь representing /y/). But with this third option these three graphemes would be representing totally different vowels.

<u>The fourth option</u> would be to use one sign *preceding* a vowel, which then signifies vowel-length consistently for all 6 vowel phonemes. Again, as with the previous option, the most logical sign to be used would be the soft-sign, as exemplified in Table 7.

Table 7 Vowels with preceding length sign

		Front						Back			
		Unr		Rnd		Unr		Rnd			
		Short	Long	Short	Long	Short	Long	Short		Long	
Close	plain	/i/ и	/iː/ ьи	/y/ уь	/yː/ ьуь			/u/ у		/uː/ ьу	
	diphthong	/ie/ иэ	/ieː/ иэ	/ye/ оь	/yeː/ оь			/uo/ уо		/uoː/ уо	
Mid		/e/ э	/eː/ ьэ					/o/ о		/oː/ ьо	
Open						/a/ а	/aː/ ьа				

This system definitely is, with some additional modifications, a real possibility. Most of the existing vowels are represented as they were until now. The three vowels that are written with a following soft-sign (the аь, оь and the уь) could still be written like that in this new system, providing backwards compatibility.

Unfortunately there is a Chechen consonant that is written with a following soft sign. The /ħ/ is written as хь. So for this option to work correctly, the way in which the /ħ/ is written should be changed. One possibility would be to write the pharyngeal fricative as a хъ. When that option is accepted there is only one potential area of conflict left. In the current system the soft sign is used as a syllable divider after the

/k/. But a new syllable can never begin with a vowel after the /k/, so there should be no problem here too. A word that is traditionally spelled as хӏаллакьеш will in the new system be spelled as хӏаллакйиэш.

A fifth option would be to use the doubling of vowels in order to write long vowels. The short vowel /a/ would then be written as a, whereas the long vowel /aː/ would be written as aa. The idea to double vowels in order to signal length is used in Latin-based orthographies, as for instance in Finnish. I have not heard of a Cyrillic-based orthography that uses vowel doubling to signal length. On the other hand, long consonants are already written by the doubling of the first letter of the consonant. So vowel doubling could be seen as an extension to that system.

In the existing orthography juxtaposition of two vowel graphemes must be handled by the reader in two different ways. When long diphthongs are written out fully, such as in уо and иэ, the vowels should be read together as forming a unit. But in other cases a glottal stop should be read between the vowels, as for example in даа (from /daʔa/ 'to eat') and хоуьйту (from /xoʔyːtu/ 'make known PRES').

The proposed vowel doubling orthography makes it necessary to abandon this dichotomy, since a word such as даа should in this proposed orthography be pronounced as /daː/ 'father'. The word /xoʔyːtu/ should then be written as хоъууьту.

Instead of using the э for all the e-sounds, I suggest to use the symbol e, and completely do away with the symbol э.

Table 8 Vowels written using doubling

		Front						Back			
		Unr		Rnd				Unr		Rnd	
		Short	Long	Short		Long		Short	Long	Short	Long
Close	plain	/i/ и	/iː/ ии	/y/ уь		/yː/ ууь				/u/ у	/uː/ уу
Close	diphthong	/ie/ ие иѐ	/ieː/ ие иé	/ye/ оь		/yeː/ оь óь				/uo/ уо уò	/uoː/ уо уó
Mid		/e/ e	/eː/ ee							/o/ o	/oː/ oo
Open						/a/ a	/aː/ aa				

In principle this fifth orthography option, the vowel doubling one, would make it possible to distinguish between short and long *diphthongs*[5]. However, the short and long diphthongs are almost (but not entirely) in complementary distribution. That is to say, the short diphthongs normally only occur in closed syllables, whereas the long diphthongs only occur in open syllables. There are a few cases where the *short* syllable occurs in *open* (non word-final) syllables. But there are no known minimal pairs, so for practical purposes I recommend that the orthography in general makes no difference between them. This is in line with the observation that the usage of graphemes should be evaluated against the functional load they carry (Seifart 2006:280).

It would be nice to have a system whereby the length of a diphthong could be indicated where that is necessary, as for instance in primers or in dictionaries. For

---

[5] One possibility would be to write the first grapheme of the diphthong twice, yielding ие, иие, оь, ооь, уо and ууо. A second possibility would be to write the second part of a long diphthong consistently with the symbol a. Thus the short diphthongs would be written as: ие, оь and уо. The long diphthongs then become as follows: иа, оа and уа.

such specific cases I suggest that accent sings be used. The short diphthongs could be written as *иѐ, ѐь, уѐ*, and their long counterparts as *иé, óь, уó*.

4.1.4. Cyrillic-based orthography proposal

If there are compelling reasons to retain the Cyrillic basis of the orthography for Chechen, then my proposal would be to make a drastic change in the orthography along the following lines.

First I would recommend a few changes in the way the consonants are written. The complete system of consonants is illustrated in Table 9. The changes with respect to the existing system are twofold:

- The glide /j/ is consistently written as й. (This implies that the three graphemes я and ю are not used in Chechen words anymore.)
- The glottal stop /ʔ/ is, whenever it is not automatically inserted in the beginning of a word, always written as a hard sign ъ.

Table 9 Cyrillic-based consonants for proposed orthography

	Noncontinuant Obstruents				Continuant Obstruents		Resonants		
							Nasal	Liquid	Glide
	Vl	Vd	Ejec	Fortis	Vl	Vd	Vd	Vd Vl	
Labial	/p/ п	/b/ б	(/p'/) пI	/p:/ пп			/m/ м		/w/ в
Dental	/t/ т	/d/ д	/t'/ тI	/t:/ тт				/l/ л	
Alveolar	/ts/ ц	/dz/ дз	/ts'/ цI	/s:/ сс	/s/ с	/z/ з	/n/ н	/r/ р /r̩/ рхI	
Palatal	/tʃ/ ч	/dʒ/ дж	/tʃ'/ чI		/ʃ/ ш	/ʒ/ ж			/j/ й
Velar	/k/ к	/g/ г	/k'/ кI	/xk/ хк	/x/ х				
Uvular	/q/ кх		/q'/ къ	/q:/ ккх		/ʁ/ гI			
Pharyngeal		/ʕ/ I			/ħ/ хь				
Glottal		/ʔ/ ъ			/h/ хI				

For the vowel system I would suggest to use the fifth option given in section 4.1.3, which is illustrated in Table 8. This option has the following advantages:

- The short vowels and most of the short diphthongs[6] are all written as they are in the existing orthography.
- The long vowels are written in a consistent way that coincides with the way consonant doubling is represented.
- Because of the low functional load the difference between long and short diphthongs is not usually made.
- For dictionaries and primers the difference between long and short diphthongs can in principle be made using accent signs.
- It remains possible to use the grapheme аь to represent the sound [æ].

There are also disadvantages to this proposal, from which I would like to mention the following:

---

[6] There is one slight difference. In the current orthography one diphthong is written as *иэ*, whereas the new orthography proposes it to be written as *ие*, since the vowel э is completely done away with.

- Automatically reading a glottal stop between two vowel graphemes is not possible anymore. Glottal stops will have to be written out fully (i.e. using the hard-sign).
- The way in which long /iː/ and /yː/ are written in the new system, using ии and ууь, radically differs from the way they are written now: ий and уьй.
- The way in which the diphthong /ye/ is represented, namely as оь, is still inconsistent with аь and уь, where the soft sign is used to signal a different vowel quality[7].

4.2. Latin-based orthography proposal

As the Chechen diaspora grows and the internet era develops, several different Latin-based orthographies have come into use. I will not review these here. Instead the proposal in this section is based on a Latin orthography that has been successfully used for the Chechen language, as well as the closely related Ingush language, among scholars for more than 10 years now. This orthography was originally developed by Johanna Nichols (Nichols 1994), and later slightly adapted. Its original purpose was to provide an orthography by which linguistic data could be easily shared between linguists. For that reason only the "lower ASCII" symbols and no diacritics were used.

4.2.1. Proposal for consonants

The way in which the consonant phonemes are represented is as follows. The sounds that more or less coincide with English are written with the same graphemes as in English. This includes writing /ʃ/ and /tʃ/ as *sh* and *ch* respectively. A logical extension to this system includes writing /ʒ/ as *zh*.

The phoneme /ʁ/ is not known as such in English. In the Latin-based orthography it is written as *gh*.

To represent the phoneme /x/ the decision was made to write it using the English *x*, which would otherwise have been left unused.

Other logical extensions to the system evolving above included writing the uvular plosive as *q* and writing the voiceless trill as *rh*.

The glide /w/ is written as a *v*, because the letter *w* is used to signal the phoneme /ʕ/, which is totally unknown to English. The related pharyngeal fricative /ħ/ is written as a combination of the symbol for the pharyngeal sound and the symbol for a glottal fricative, i.e. as *hw*. There is one exception: whenever the pharyngeal fricative /ħ/ follows a voiceless consonant, it is written as a single *w*. See the examples (9) ,(10) and (11). The grapheme *w* represents a phonological /ʕ/ when following a voiced consonant, as illustrated in example (12).

---

[7] These inconsistencies could only be avoided if it were decided to make even more radical changes. In particular the way in which the /y/ is written should be changed for instance to э (or ы). The long vowel /yː/ can then be written as ээ (or ыы). Similarly, the diphthongs /ye/ and /yeː/ could then be written as эе and эа (or as ые and ыа). But such changes do not look very nice, and make a switch even more difficult for those who are currently literate in Chechen.

9) *hwaasha*   /ħaːʃa/   'guest'

10) *cwoogal*   /tsħoːgal/   'fox'

11) *dieshwalxie* /dieːʃħalxieː/ 'foreword'

12) *bweestie*   /bʕeːstieː/   'spring'

The glottal stop is consistently written as a single quotation mark ', which is also used as a modification after consonants *p, t, c, ch, k* and *q* to represent the ejective consonants. Unlike with a Cyrillic based orthography, there are no inconsistencies for the writing of the k-sounds (see section 3.4.1).

Table 10 Latin-based consonants for proposed orthography

	Noncontinuant Obstruents				Continuant Obstruents		Nasal	Liquid	Glide
	Vl	Vd	Ejec	Fortis	Vl	Vd	Vd	Vl	
Labial	/p/ p	/b/ b	(/p'/) p'	/pː/ pp			/m/ m		/w/ v
Dental	/t/ t	/d/ d	/t'/ t'	/tː/ tt				/l/ l	
Alveolar	/ts/ c	/dz/ dz	/ts'/ c'	/sː/ ss	/s/ s	/z/ z	/n/ n	/r/ r /r̩/ rh	
Palatal	/tʃ/ ch	/dʒ/ dzh	/tʃ'/ ch'		/ʃ/ sh	/ʒ/ zh			/j/ j
Velar	/k/ k	/g/ g	/k'/ k'	/xk/ xk	/x/ x				
Uvular	/q/ q		/q'/ q'	/qː/ qq	/ʁ/ gh				
Pharyngeal		/ʕ/ w			hw, /ħ/ w				
Glottal		/ʔ/ '			/h/ h				

4.2.2. Proposal for vowels

In this proposal the short and long vowels are written much like in Finnish. The written vowels correspond to the phonemes as written with IPA symbols, as illustrated in Table 11.

The representation of the short diphthongs /ie/, /uo/ and /ye/ coincides with their IPA representation, and is *ie*, *uo* and *ye* respectively. For the short diphthongs the stress is on the second part of the diphthong, and by using this orthography, the pronunciation is reflected accurately. The representation of the corresponding long diphthongs /ieː/, /uoː/ and /yeː/ consistently uses the *a* as second letter in the grapheme, resulting in graphemes *ia*, *ua* and *ya*. In long diphthongs the accent is usually on the first part of the diphthong. The pronunciation of the first part is reflected accurately by the proposed orthography. The unaccented second part is a schwa kind of sound, which is accurately represented with the *a*, since the unaccented short *a* in Chechen actually is a schwa.

The representation of the long and short vowel diphthongs is an extension to the orthography used by Nichols (2007). The latter orthography does not distinguish between short and long diphthongs[8]. This is a wise decision, although, strictly

---

[8] I leave the diphthong represented by Nichols as *ia* out of the discussion here. This diphthong is a sound, whose phonemic nature I have not investigated.

speaking, the long and short diphthongs are not in complementary distribution. That is to say, there are a few near minimal pairs of these diphthongs in open syllables. For example the word /sieda/ 'star' contains a short /ie/ diphthong in the first open syllable, whereas the word /sieːga/ 'to radiate' contains a long /ieː/ diphthong in a similar environment. But generally short diphthongs occur in closed syllables, while long dipthongs occur in open ones. For practical reasons I suggest to write the short and long diphthongs identically in general, as is done in Nichols' orthography. But I would suggest an extension to this system, in order to allow the short/long distinction to be made when an author feels this is important (e.g. for pedagogical reasons). I propose to use diacritics on the second vowel of the graphemes to signal the length of the diphthong. A falling accent sign (like the accent grave in French as in *là* 'there') can be placed on the diphthong showing shortness, resulting in *iè*, *uò* and *yè*. Similarly, if it needs to be shown that the diphthong is the phonemic long one, then a rising accent sign (like the French accent aigu as in *malgré* 'despite') can be placed, resulting in *ié*, *uó* and *yé*. The rising accent sign is also used to denote phonemic long vowels in Hungarian.

Table 11 Latin-based vowels for proposed orthography

	Front				Back				Diphthong	
	Unr		Rnd		Unr		Rnd			
	Short	Long	Short	Long	Short	Long	Short	Long		
Close	/i/ i	/iː/ ii	/y/ y	/yː/ yy			/u/ u	/uː/ uu	/ui/ ui	
	/ie/ ie iè	/ieː/ ie ié	/ye/ ye yè	/yeː/ ye yé			/uo/ uo uò	/uoː/ uo uó	/uoi/ uoi	/ou/ ou
Mid	/e/ e	/eː/ ee					/o/ o	/oː/ oo		/ey/ eu
Open					/a/ a /aː/ aa				/ai/ ai	

4.2.3. Discussion of the Latin-based orthography proposal

There are a number of advantages of the proposed Latin orthography for Chechen over the currently used Cyrillic one:
- The proposed main orthography makes use of Latin characters that are standardly available on any computer. No diacritics are needed for the main orthography. The extension makes use of diacritic letters that are also standard available on modern computers.
- Using the Latin orthography makes computer communication very easy. It also means that the largest selection of fonts is available to any user. Development of computer tools to support the Chechen language does not need to make use of non-standard characters.
- A Latin based orthography makes it much easier for second-generation Chechen speakers who find themselves as refugees in Western countries to learn to read and write their mother tongue.
- There are no inconsistency problems with the k-sounds.
- All vowel phonemes are representable in a consistent and logical way.
- Any Chechen who is literate in the Latin based orthography will have less problems learning a Western language, as he already knows the basics of the orthography.

Even this Latin-based orthography has its (potential) problems. The following ones can be identified:

- If Chechen children learn to read and write their mother tongue in the first grade, they will do so using a Latin based orthography. When in later grades they start learning Russian, they need to learn another orthography (the Cyrillic one). Learning a different orthography is a large step. On the other hand – it does make the vast difference between the two languages clearer for early learners.
- The usage of *w* and *hw* for the epiglottal stop and the pharyngeal fricative requires good teaching for those who are used to the current Cyrillic orthography.
- Any *hw* sound following a voiceless consonant is written with one grapheme, the *w*, resulting for instance in *cw*, *chw*, *sw* and *shw*. This contrasts with the voiced epiglottal consonant written as *w*. When this sound is preceded by a consonant, that consonant can only be a voiced one, as for instance in *dw*, *bw*, *zw*, *zhw*, *mw*, *nw*. This distinction requires careful attention in primers.
- Whatever current Chechen literature there is would have to be converted into the Latin orthography. The conversion would be semi-automatic, because for many words a choice needs to be made as to the phonemic nature of the vowels.
- The apostrophe that is used to show the glottal stop is not a standard *letter* in Latin-based orthographies. This offers some difficulties in word processors, where it may be mistaken for a direct-speech single quote, when it comes word-finally as fo instance in /qaʔ/, 'good news', which would be written as *qa'*. A doubled glottal stop is graphically indiscernable from a double quote, as for instance in the word /daʔːalts/, 'until eating', that would be written in the Latin orthography as *da"alc* (double glottal stops don't occur word-finally however).

With the Latin-based orthography proposal Chechen is supposed to be written almost completely phonemically. The orthography does not require one to make the phonemic difference between long and short diphthongs (writing accent signs to indicate that the diphthong is short or long is optional). But note that in closed syllables the difference between short and long diphthongs is neutralized, and in some cases it is not possible to find out whether the underlying form of the diphthong is short or long, as for instance in the word *biexk* 'guilt'. Since no obligatory choice has to be made between a grapheme representing a phonemically short diphthong and a grapheme for the long diphthong, the potential problem of which choice to make is conveniently avoided.

5.   Words illustrating the writing of the Chechen vowel phonemes

In order to illustrate the impact that the proposed orthographies would be making on the way words are written, in particular with respect to the vowels, the list of examples below is included. For each vowel phoneme one or two words are given in the existing orthography, the proposed Cyrillic orthography and the proposed Latin orthography.

/i/	/χi/	хи	хи	xi	'water'
	/jiʃa/	йиша	йиша	jisha	'sister'
	/xir/	хир	хир	xir	'will be'

/iː/	/iːtsira/	ийцира	иицира	iicira	'bought'
	/jiːʃa/	йийша	йииша	jiisha	'to sleep'
	/piːl/	пийл	пиил	piil	'kick'
/ie/	/biexk/	бехк	биехк	biexk	'guilt'
	/iexna/	эхна	иехна	iexna	'went/flew'
	/sieda/	седа	сиѐда	sièda	'star'
/ieː/	/dieː/	де, диэ	дие	die	'do!'
	/dieːʃ/	деш	диеш	diesh	'doing'
	/sieːga/	сега, сēга	сиѓга	siéga	'to radiate'
/e/	/de/	де	де	de	'day'
	/xezira/	хезира	хезира	xezira	'heard'
	/kʼel/	кӀел	кӀел	k'el	'under'
/eː/	/xeːnaħ/	хенахь	хеенахь	xeenahw	'at the time'
	/dæltʃa/	даьлча	даьлча	daelcha	'when went'
/y/	/xylu/	хуьлу	хуьлу	xylu	'happens'
	/jyxk/	юьхк	йуьхк	jyxk	'stump of wood'
/yː/	/syːrieː/	суьйре	сууьриа	syyria	'evening'
	/nyːr/	нуьйр	нууьр	nyyr	'saddle'
/ye/	/yexu/	оьху	о̀ьху	yèxu	'comes/flies'
	/tyeʔna/	тоьъна	тоьъна	tye'na	'has sufficed'
/yeː/	/tʃyeː/	чоь	чоь	chye	'inside'
	/syeːga/	соьга	соьга	syega	'to me'
	/tyeːxna/	тоьхна	тоьхна	tyexna	'has hit'
/a/	/xa/	ха	ха	xa	'guard'
	/xabar/	хабар	хабар	xabar	'talk, rumour'
	/ʃad/	шад	шад	shad	'whip'
/aː/	/daː/	да	даа	daa	'father'
	/xaːda/	хада	хаада	xaada	'to cut'
	/jaːlxitːa/	ялхитта	йаалхитта	jaalxitta	'sixteen'
/u/	/du/	ду	ду	du	'is'
	/tʼexula/	тӀехула	тӀехула	t'exula	'above'
	/nuskal/	нускал	нускал	nuskal	'bride'
/uː/	/luː/	лу	луу	luu	'kind of deer'
	/suːna/	суна	сууна	suuna	'to me'
	/guːr/	гур	гуур	guur	'trap'
/uo/	/tuoʔa/	тоа	туѐъа	tuò'a	'to suffice'
	/xuox/	хох	хуох	xuox	'union'
/uoː/	/suoː/	суо	суо	suo	'I (emphatic)'
	/duoːʁa/	догӀа	дуогӀа	duogha	'to build, plant'
	/xieːxuoːʃna/	хехошна	хиехуошна	xiexuoshna	'to the guards'
/o/	/txo/	тхо	тхо	txo	'we (excl)'
	/doʁa/	догӀа	догӀа	dogha	'rain'
	/noxtʃiː/	нохчий	нохчии	noxchii	'Chechens'
/oː/	/soːnieː/	соне	соодие	soonia	'into the corner'
	/doːguʃieħ/	доггушехь	доогушиехь	dooggushiehw	when burns'
/ui/	/dui/	дуй	дуй	dui	'is?'
/uoi/	/xieːxuoi/	хехой	хиехуой	xiexuoi	'guards'
/ai/	/sai/	сай	сай	sai	'deer'
/ou/	/kou/	ков	ков	kou	'gate'
/ey/	/deyzig/	девзиг	девзиг	deuzig	'broom'

The proposed orthographies are further illustrated in the transcription of a sample text in the Appendix in section 10.

6. Conclusions and discussion

A review of the Chechen sound system has revealed that, on the whole, the consonants are treated fairly in the current Chechen Cyrillic based orthography. A notable problem exists with the usage of the я, ю and the е where words start with a noun-class prefix /j/. Several possibilities to improve the way in which consonants are written in Chechen were discussed. For a Cyrillic-based orthography the conclusion is that the symbols я and ю should not be used anymore and that the noun-class prefix /j/ should consistently be written as й. No feasible solution for the inconsistency in writing the k-sounds (in particular the /q'/) has been found within the Cyrillic orthography.

On the other hand a Latin-based orthography does not contain inconsistencies as far as the k-sounds are concerned. Nor is there any problem in faithfully and consistently representing the noun-class prefixes.

A thorough review of the Chechen vowel phonemes made it clear that the currently used Cyrillic orthography lacks the means to make the differentiations necessary. Several options to radically change the way in which vowels are written were reviewed. Within the possibilities offered by the Cyrillic orthography the best option seems to be to write long vowels by repeating the first part of the grapheme representing the short vowel. The grapheme э should not be used any longer, but replaced with е consistently, even word-initial. Short and long diphthongs should be written out as ие, уо and оь. When it is necessary to distinguish between short and long diphthongs, such as in dictionaries or in primers, they can be written as иѐ, о̀ь, уо̀ or иé, óь, уó respectively.

A separate proposal was made how to write all the different vowel phonemes in a Chechen orthography based on the Latin script. Short vowels are then written as *i, e, a, u, o, y* and their long counterparts as *ii, ee, aa, uu, oo, yy*. The diphthongs are in general written as *ie, uo* and *ye*. When it is necessary to distinguish between short and long diphthongs they can be written as *iè, uò, yè* or *ié, uó, yé* respectively.

Switching from the currently used Cyrillic based orthography to the proposed Latin based orthography has a number of advantages:
- The problem with writing the k-sounds is eliminated.
- All short vowels are written with one character (no diacritics are needed for the short vowels).
- Latin characters with diacritics provide a good and logical way to differentiate between short and long diphthongs wherever needed.
- Switching to a totally different basis for the orthography provides a "clean sweep" for everyone.
- This orthography is easier for children of refugees living in Western countries.
- Anyone anywhere will be able to write with this Latin based orthography, since it only makes use of characters that are standardly available on any computer. This is "standard" in the sense that any Windows-based computer will have all the graphemes for the orthography, as well as the vowels containing diacritics to optionally distinguish between short and long diphthongs, available under the "English-International" keyboard.

There are also disadvantages to starting to use a Latin-based orthography, listed here:

- Even though the percentage of Chechen people who are literate in their own language is very small, a switch to any new orthography will be difficult for them. And a switch to a Latin-based orthography will be even more difficult to them.
- With the advocated new orthography Chechen children living in their homeland would in the first grade learn to read and write their mother tongue using the Latin script. But in the next grades they would have to learn the Cyrillic script in order to read and write Russian. This might not be such a big disadvantage, since the languages are so vastly different anyway. For Chechen children living outside the homeland the usage of a Latin script actually would be an advantage.

Whatever switch is going to be made, the impact of an orthography reform will be felt through the whole Chechen society. Scientists are needed to determine how the phonemic layout of words is. Dictionaries will be needed, in order to know how to properly write a word with a certain meaning. Primers will need to be made based on the new orthography. Existing literature is not going to be transformed into the new orthography just like that.

It is exactly because of the large impact of making changes, that I suggest it would be good to adopt an orthography that will once and for all be sufficient for the Chechen language from an academic point of view, and so consistent, that learning to read and write in it is as easy as possible for anyone.

I hope that the considerations, options and proposals given in this paper are of use to the committee that is currently considering an orthography and spelling reform for the Chechen language.

7. Acknowledgments

8. Affiliations

Erwin R. Komen
Radboud University   SIL-International
CLS-ETC   Erwin_Komen@SIL.org
Box 9103
6500 HD Netherlands
E.Komen@Let.RU.nl

9. References

Dzhamalxanov Zaindi D., Machigov M.Yu. 1972. *Noxchiin Muott.* Part 1: *Leksikologi, fonetika, morfologi.* Grozny.

Dzhamalxanov Zaindi D., Aliroev I.Yu. 1992. *Noxchiin mettan niisajaazdaran dosham.* Grozny.

Timaev Vaaxa. 2007. Interview on radio: *Chechnya Svobodnaya.* (June 18th, 2007.)

Komen Erwin R. 1993. *Chechen Phonology Interim Report.* Manuscript.

Komen Erwin R. 1996. *Interaction between phonology, orthography and literacy in the Chechen vowel system*, Paper presented at the Societas Caucasologica in Leiden, Netherlands.

Komen Erwin R. 2007. *Chechen Vowel Inventory.* To appear in: Pepijn Hendriks, Frank Landsbergen, Mika Poss & Jenneke van der Wal (eds.), *Leiden Papers in Linguistics.* Vol. 4, issue 2.

Maciev A.G. 1961. *Чеченско-Русский словарь.* Moscow.

Nichols, Johanna. 1994. *Chechen.* In: Rieks Smeets (ed.) *The indigenous languages of the Caucasus*, vol. 4 *The North East Caucasian languages II.* Delmar, Caravan books.

Nichols, Johanna. 1997. *Chechen Phonology.* Chapter 47 in: Alan S. Kaye (ed.) *Phonologies of Asia and Africa.* Vol. 2. Eisenbrauns. Winona Lake, USA. 941-971.

Nichols, Johanna. 2007. *The Chechen sound system, in the Latin transcription used here.* http://socrates.berkeley.edu/%7Echechen/Ch_sdcht_lat.htm. (August 27th, 2007.)

Seifart, Frank. 2006. *Orthography development.* In: Jost Gippert, Nikolaus P. Himmelmann, Ulrike Mosel (eds.) *Essentials of language documentation.* Mouton de Gruyter. Berlin. 275-299.

Xamidova, Zulai. 2007. *Noxchiin tyyranash.* http://www.zhaina.com. (August 28th, 2007.)

10. Appendix

In this appendix the first part of one Chechen fairy tale is reproduced in different orthographies. The story is called "The bird of crane and the fox" (Xamidova 2007).

10.1. Existing orthography

ГӀарагӀули а, цхьогал а

Юххехь Ӏар-дахар нисделла гӀарагӀулин а, цхьогалан а. Дика лулахой а хилла Ӏаш хилла уьш. ГӀарагӀули, шена бен а бина, хӀоаш тӀе хиъна. Хан-зама яьлча, кӀорнеш даьхна цо. Цхьогало а кӀезий дина. ХӀор Ӏуьйранна гӀарагӀули сатоссуш ижуӀ лаха йоьдура, ткъа цхьогал цӀахь дуьсура, цуьнан а, шен а берашка а хьожуш, дан дезарг а деш.

ГӀарагӀули дӀа ма яхханехь, цхьогало гӀарагӀулин кӀорни схьа а лаьцна, корта а хьовзийна, ур-атталла цхьа даьӀахк а ца юьтуш, дӀакхаьллира. ГӀарагӀули цӀа еъча, доьхна хьаьвзира цхьогал:

— Ва, гӀарагӀули, ма чӀогӀа вон деъна-кх вайга! — элира цо.

— ХӀун хилла? — хаьттина гӀарагӀулино.

— Вайн цхьа бер дайна, мича дахана а ца хууш. Аса массо а меттахь лийхира и, суна цхьаннахьа а ца карийра.

— Хьенан кӀорни дайна? — хаьттира гӀарагӀулино.

— Ой, вайниг дайна-кх, вайниг, — элира цхьогало.

— Хьайниг дайна я сайниг дайна?

— Еха ворта а, еха настарш а йолуш дерг дайна.

Кхийтира гӀарагӀули шен кӀорни дайний. Цхьогал дукха чогӀа холчухӀоьттина хьаьвзича, гӀарагӀулино элира:

— Са ма гаттаде, хилларг хилла даьлла, иштта яздина хилла-кх вайн кӀорнина.

— И бакъ ду, — тӀетайра цхьогал, — иштта кхоллам хилла-кх цуьнан.

10.2. Accented Cyrillic orthography

See section 4.1.3, the first option.

ГӀарагӀули а, цхьо́гал а

Йуххêхь Ӏар-да́хар нисделла гӀа̀рагӀулин а, цхьо́галан а. Дика лулахôй а хилла Ӏаш хилла уьш. ГӀарагӀули, шêна бêн а бина, хӀоъаш тӀе хиъна. Ха́н-за́ма яьлча, кӀôрнêш даьхна цô. Цхьо́галô а кӀезий дина. ХӀо́ра Ӏу́ьранна гӀарагӀули сатоссуш ижу́ лаха йо́ьдура, ткъа цхьо́гал цӀахь дуьсура, цуьнан а, шêн а бе́рашка а хьожуш, дан дêзарг а деш.

ГӀарагӀули дӀа ма йа̀хханêхь, цхьо́галô гӀарагӀулин кӀôрни схьа а лаьцна, кôрта а хьовзиина, ур-атталла цхьа даьӀахк а ца йуьтуш, дӀа́кхаьллира. ГӀарагӀули цӀа йеъча, доьхна хьаьвзира цхьо́гал:

— Ва́, гӀарагӀули, ма чӀо́гӀа вôн деъна-кх вайга! — е́лира цô.

— ХӀун хилла? — хаьттина гӀарагӀулинô.

— Вайн цхьа бе́р дайна, мича да̀хана а ца хуъуш. Аса массô а меттахь лийхира и, су́на цхьаннахьа а ца карийра.

— Хье́нан кӀôрни дайна? — хаьттира гӀарагӀулинô.

— Ой, вайниг дайна-кх, вайниг, — е́лира цхьо́галô.

— Хьайниг дайна йа сайниг дайна?

— Йе́ха вôрта а, йе́ха настарш а йолуш дерг дайна.

Кхийтира гӀарагӀули шêн кӀôрни дайний. Цхьо́гал дукха чӀо́гӀа холчухӀоьттина хьаьвзича, гӀарагӀулинô е́лира:

— Са ма гаттадê, хилларг хилла даьлла, иштта йа́здина хилла-кх вайн кӀôрнина.

— И бакъ ду, — тӀетайра цхьо́гал, — иштта кхôллам хилла-кх цуьнан.

10.3. New Cyrillic orthography

ГӀарагӀули а, цхьоогал а

Йуххиехь Ӏар-даахар нисделла гӀарагӀулин а, цхьоогалан а. Дика лулахуой а хилла Ӏаш хилла уьш. ГӀарагӀули, шиана биен а бина, хӀоъаш тӀе хиъна. Хаан-заама яьлча, кӀуорниеш даьхна цуо. Цхьоогалуа а кӀезии дина. ХӀоора Ӏууьранна гӀарагӀули сатоссуш ижуу лаха йоьдура, ткъа цхьоогал цӀахь дуьсура, цуьнан а, шиен а беерашка а хьожуш, дан диезарг а диеш.

ГӀарагӀули дӀа ма йаххание хь, цхьоогалуо гӀарагӀулин кӀуорни схьа а лаьцна, куорта а хьовзиина, ур-атталла цхьа даьӀахк а ца йуьтуш, дӀаакхаьллира. ГӀарагӀули цӀа йеъча, доьхна хьаьвзира цхьоогал:

— Ва, гӀарагӀули, ма чӀоогӀа вуон деъна-кх вайга! — еелира цуо.

— ХӀун хилла? — хаьттина гӀарагӀулинуо.

— Вайн цхьа беер дайна, мича дахана а ца хуъуш. Аса массуо а меттахь лийхира и, сууна цхьаннахьа а ца карийра.

— Хьеенан кӀуорни дайна? — хаьттира гӀарагӀулинуо.

— Ой, вайниг дайна-кх, вайниг, — еелира цхьоогалуо.

— Хьайниг дайна йа сайниг дайна?

— Йееха вуорта а, йееха настарш а йолуш дерг дайна.

Кхиитира гӀарагӀули шиен кӀуорни дайнии. Цхьоогал дукха чоогӀа холчухӀоьттина хьаьвзича, гӀарагӀулинуо еелира:

— Са ма гаттадие, хилларг хилла даьлла, иштта йааздина хилла-кх вайн кӀуорнина.

— И бакъ ду, — тӀетайра цхьоогал, — иштта кхуоллам хилла-кх цуьнан.

10.4. New Latin orthography

Gharaghuli a, cwoogal a

Juxxiehw war-daaxar nisdella gharaghulin a, cwoogalan a. Dika lulaxuoj a xilla wash xilla ysh. Gharaghuli, shiena bien a bina, ho'ash t'e xi'na. Xaan-zaama jaelcha, k'orniesh daexna cuo. Cwoogaluo a k'ezii dina. Hoora wyyranna gharaghuli satossush izhuu laxa jyedura, tq'a cwoogal c'ahw dysura, cynan a, shien a beerashka a hwozhush, dan diezarg a diesh.

Gharaghuli dwaa ma jaxxaniehw, cwoogaluo gharaghulin k'uorni swa a laecna, kuorta a hwovziina, ur-attalla cwa daewaxk a ca jytush, dwaaqaellira. Gharaghuli c'a je'cha, dyexna hwaevzira cwoogal:

— Va, gharaghuli, ma ch'oogha vuon de'na-q vaiga! — eelira cuo.

— Hun xilla? — xaettina gharaghulinuo.

— Vain cwa beer daina, micha daxana a ca xu'ush. Asa massuo a mettahw liixira i, suuna cwannahwa a ca kariira.

— Hweenan k'uorni daina? — xaettira gharaghulinuo.

— Oi, vainig daina-q, vainig, — eelira cwoogaluo.

— Hwainig daina ja sainig daina?

— Jeexa vuorta a, jeexa nastarsh a jolush derg daina.

Qiitira gharaghuli shien k'uorni dainii. Cwoogal duqa ch'oogha xolchuhyettina hwaevzicha, gharaghulinuo eelira:

— Sa ma gaattadie, xillarg xilla daella, ishtta jaazdina xilla-q vain k'uornina.

— I baq' du, — t'etaira cwoogal, — ishtta quollam xilla-q cynan.