

Phonological and Orthographical Information in Dictionaries: The Case of Pröhle's Karachay Glossary and its Successors

Steve Seegmiller

It has often been pointed out that a dictionary is more than a book where one can look up the meanings of words. In this paper, I would like to focus on two other important uses of dictionaries: first, dictionaries are valuable sources of data for researchers studying a language; and second, dictionaries are probably the primary tool for establishing orthographical norms, especially for new writing systems. I will use as a case study two works on Karachay, an early glossary and a more recent large dictionary.

Karachay is a Turkic language spoken in the Northern Caucasus. In Soviet publications it is usually referred to as Karachay-Balkar for its two main varieties. However, I will refer to the language simply as Karachay since in my own research on the language I have dealt only with speakers of Karachay and I have no first-hand knowledge of Balkar.

This is a particularly appropriate time and place to deal with this topic because it was in this city, nearly eighty years ago, that the first glossary of Karachay appeared in print. I am referring to Wilhelm Pröhle's "Karatschajisches Wörterverzeichnis", (Pröhle 1909b). As I will try to show, Pröhle's work continues to be influential to this day, even beyond its original intended purpose.

My interest in this topic comes from my own work on Karachay. For the past several years, I have been working with some colleagues on a dictionary of Karachay, Turkish and English. As I began to examine the published material on Karachay, I quickly located Pröhle's "Karatschajisches Wörterverzeichnis" and the accompanying "Karatschajische Studien" (Pröhle 1909a) since they remain the standard reference works on Karachay in the West. I further discovered that virtually all publications on Karachay since Pröhle's two works are from the Soviet Union, including a 1965 Russian-Karachay dictionary by Sujunčev and Urusbiev entitled *РУССКО-КАРАЧАЕВО-БАЛКАРСКИЙ СЛОВАРЬ*. In examining the latter work, I was puzzled to find that it incorporated an overdifferentiated orthography, as do all other Soviet works on Karachay that I have seen. What I mean by an overdifferentiated orthography is one in which one or more nonphonemic differences are spelled differently. It is surprising to find such overdifferentiation in a writing system that is scarcely fifty years old, the Cyrillic writing system for Karachay having been introduced in 1938 or 1939. One expects a newly devised writing system to be as close to the ideal as possible. By an "ideal" writing system I mean one that is based on the phonemic system of the language, that is, one which spells different phonemes differently but disregards non-phonemic differences. The reason that a phonemically-based writing system is desirable in most cases is that it is the easiest kind for speakers of a language to master. A phonemic writing system matches their intuitions about their language and therefore learning to write is very simple. If a writing system contains some non-phonemic distinctions, speakers of the language will have trouble hearing them and consequently the spellings of many words will have to be learned through rote memorization. Obviously, if the goal of

devising a new writing system is to promote literacy among the speakers of the language, then an ideal writing system will make this task as easy as possible. Equally obvious is the role of dictionaries in establishing literacy: if people want to know how to spell a word, they look it up in a dictionary. Therefore, a dictionary intended to help promote literacy should incorporate an ideal writing system.

Examples of "ideal" or near-ideal writing systems in this sense include the orthographical system of Turkish, which makes use of the Roman alphabet and was introduced about sixty years ago; and the present Russian writing system, which is the result of an orthographical reform introduced following the Revolution of 1918. There are many other writing systems in use around the world which approach this ideal, and their principles are well understood. It is therefore surprising to find a writing system devised in this century which deviates from the ideal in such a clear way.

In order to understand the problem with the Karachay writing system, we must first examine the phonological system of Karachay. The vowel and consonant phonemes of Karachay are presented in Tables 1 and 2.

TABLE 1
KARACHAY VOWEL PHONEMES

	<i>Front</i>		<i>Back</i>	
	<i>Rounded</i>	<i>Unrounded</i>	<i>Rounded</i>	<i>Unrounded</i>
High	ü	i	u	ĩ
Low	ö	e	o	a

TABLE 2
KARACHAY CONSONANT PHONEMES

	<i>Labial</i>	<i>Dental</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
STOPS	b p	d t		g k	
FRICATIVES	(f)	z s	š	x	h
AFFRICATES		(c)	j č		
NASALS	m	n		ŋ	
LIQUIDS		l r			
GLIDES	w		j		

(Note: the sounds enclosed in parentheses occur only in borrowed words.)

One notable feature of Karachay phonology is that the two velar stop phonemes—/k/ and /g/—each has an unusual set of allophones. /k/ is pronounced as a velar stop /k/ next to front vowels and as a voiceless uvular stop [q] next to back vowels. (g) is pronounced as a voiced velar stop [g] next to front vowels, as a voiced velar fricative [ɣ] next to back vowels, and, depending on the variety, as either a voiced uvular stop [G] or a voiceless velar fricative [x] between a stop and a back vowel.

Here are some examples of the distinctions described above, in phonemic and phonetic transcription.

Phonemic Transcription

Phonetic Transcription

(1) *Allophones of /k/*

(a) /kelir/ 'he comes'	[kelir]
(b) /eki/ 'two'	[eki]
(c) /bilek/ 'arm'	[bilek]
(d) /kabar/ 'he bites'	[qabar]
(e) /sakal/ 'beard'	[saqal]
(f) /ačik/ 'open'	[ačiq]

(2) *Allophones of /g/*

(a) /günax/ 'sin'	[günax]
(b) /egeč/ 'sister'	[egeč]
(c) /baga/ 'price, value'	[baɣa]
(d) /agač/ 'tree, wood'	[aɣač]
(e) /kapgan/ 'bitten'	[qapxan] or [qapGan]

Notice in examples (1a) through (c) that a /k/ next to a front vowel is pronounced as /k/ while in (1d) through (f) a /k/ next to a back vowel is pronounced as /q/. In (2a) and (b) we see that a /g/ next to a front vowel is pronounced /g/; in (2c) and (d) that a /g/ next to a back vowel is pronounced as [ɣ], except that a /g/ between a stop and a back vowel is pronounced as [x] by some speakers of Karachay and as [G] by others, as we see in (2e). It is important to emphasize once again that these sounds are all allophones of the phonemes /g/ and /k/, and therefore should not be spelled differently in an ideal orthography.

Pröhle's "Karatschajisches Wörterverzeichnis" (Pröhle 1909b) uses the following set of symbols for transcribing Karachay, in alphabetical order: a, b, c, č, d, e, f, g, γ, h, x, i, j, k, q, l, l̄, m, n, ŋ, o, ö, p, r, s, š, t, u, ü, w, z, ž. Notice that the eight vowels are all transcribed as in the vowel chart above. Pröhle's ž is equivalent to [j]. For our purposes, the thing to note in this list is that Pröhle records three non-phonemic distinctions in his transcription system: γ, q, and l̄. The last of these, l̄, is described by Pröhle as "schwach palatisiert," or lightly palatalized. The other two, γ and q, have already been mentioned as allophones of /g/ and /k/, respectively. Pröhle records no instances of [G]. This sound occurs in the speech of some, but not all, speakers of Karachay, and I assume that none of Pröhle's informants used this sound. In cases where my informants used [G], Pröhle records [x], which is an independent phoneme in addition to being an allophone of /g/.

Let us now examine Pröhle's transcriptions of the words given in (1) and (2) above.

(3) *Words with /k/*

- (a) kelir
- (b) eki
- (c) bilek
- (d) qabar
- (e) saqal
- (f) ačiq

- (4) *Words with /g/*
 (a) gūnax
 (b) egeč
 (c) baγa
 (d) aγac
 (e) qarxan

It is clear that Pröhle has transcribed the velar sounds phonetically rather than phonemically. So we see that the very first publication on Karachay records words in an overdifferentiated spelling system.

During the early 1920's the Soviet government began to devise alphabets for the unwritten languages of the Soviet Union. For Karachay, as for the other Turkic languages that did not already have writing systems, the first writing system used was based on the Arabic syllabary in use at the time in Turkey. In the late 1920's, after Turkey had discarded its Arabic script in favor of the Roman alphabet, the Soviet government introduced a Roman alphabet for Karachay and several other languages. Then finally in 1938 or 1939, with a desire to establish uniformity, all Soviet languages except those like Georgian and Armenian that already had well-established writing systems were written in a modified Cyrillic alphabet. It is this alphabet in which all Soviet publications on Karachay have been published since 1939.

For both practical and theoretical reasons, there has been a great deal of interest in Karachay and other minority languages of the Soviet Union, and consequently the number of Soviet publications on Karachay is quite extensive. These include grammars and grammatical sketches in both Russian and Karachay, teaching materials, literary texts, etc. (A selected bibliography will be found at the end of the paper). It was not until 1965, however, that a new dictionary of Karachay appeared, this one a Russian-Karachay dictionary (Sujunčev and Urusbiev 1965) containing nearly 35,000 words (compared to Pröhle's 1500 words or so). I will discuss only this last-mentioned work, although the transcription system used there is typical of virtually all Soviet works on Karachay published in the Cyrillic alphabet. Here are the words from (1) and (2) as they appear in the Russian-Karachay dictionary.

- (5) (a) келир
 (b) еки
 (c) билек
 (d) къабар
 (e) сакъал
 (f) ачыкъ
- (6) (a) гюнах
 (b) егеч
 (c) багъа
 (d) агъач
 (e) къабхан

The Cyrillic alphabet for Karachay has clearly retained the overdifferentiation that we saw in Pröhle's transcription. The "hard sign" ъ is used as a diacritic to distinguish [k] from [q], the latter spelled къ; and similarly the hard sign distinguishes [g] from [ɣ], the latter spelled гъ. (Like Pröhle's glossary, the Russian-Karachay

dictionary contains no instances of [G].) Once again, we find the velars and uvulars written phonetically rather than phonemically.

Thus we see that in both Pröhle's Karachay glossary and the Russian-Karachay dictionary an overdifferentiated writing system is being used. I will now try to show why this writing system is the appropriate one in Pröhle's case, but not in the case of the Soviet dictionary.

The two works in question clearly had different purposes. Pröhle's glossary was published in 1909, at a time when no writing system existed for Karachay and therefore no speakers of Karachay could have been literate in their own language. Thus Pröhle's glossary was obviously intended for an audience of researchers, in particular Turkologists and those interested in the Uralo-Altaic languages. We might term Pröhle's glossary a *scholarly* dictionary. In scholarly dictionaries the most appropriate choice in selecting a transcription system is the one which provides the maximum amount of potentially relevant information to scholars. In selecting a transcription system for a scholarly dictionary, a lexicographer might justifiably adopt a fairly narrow phonetic transcription. As a matter of fact, in his "Karatschajische Studien", another work obviously intended for scholars, Pröhle recorded even more phonetic detail than in the glossary, including aspirated stops, which also are not phonemic. In scholarly dictionaries, questions as to what kind of writing system would be easiest for a speaker of the language to learn are usually not relevant. In this context, Pröhle's choice of an overdifferentiated, phonetic transcription system was the appropriate one.

On the other hand, the Soviet dictionary of Russian and Karachay was intended for an entirely different audience, namely speakers of Karachay. It is clearly normative in that a great many Karachay words, especially those dealing with science, technology, government, literature, education, and other such topics are borrowings from Russian. I would estimate that well over half of the words in the dictionary are Russian. It is obvious, then, that this dictionary is intended to establish vocabulary norms for Karachay, in part by providing words from Russian for the things and concepts for which Karachay doesn't already have words. The Russian-Karachay dictionary will also inevitably serve as the primary orthographical reference work for the language since it is the only large dictionary of Karachay available. We might call this type of dictionary a *practical* dictionary, since it is intended for use by the speakers of the language. In a practical dictionary a phonemic writing system is the most suitable choice. The orthographical overdifferentiations in the Russian-Karachay dictionary will unfortunately cause unnecessary problems for speakers of Karachay trying to learn to write, as I have already confirmed with my Karachay colleagues. They find it almost impossible to distinguish [k] from [q] or [g] from [G] and [ɣ]. Of course, speakers of Karachay will eventually learn to spell, just as countless generations of English- and French-speaking children have mastered much more difficult orthographies, but the process of learning to write has been made more difficult than it needs to be.

It is interesting to speculate on why the Cyrillic orthography for Karachay contains these overdifferentiations. It is certainly not the case that Soviet linguists are naive about phonology or writing, for it was, after all, in Nineteenth Century Russia that the first use of the term phoneme, and the first discussion of the phonemic principle, appeared in print. Nor can it be that the authors of the Russian-Karachay dictionary did not have a clear purpose in mind for that work, since it is

so clearly and intentionally normative. My conclusion is that the overdifferentiations that are found in Soviet publications on Karachay are the result of tradition, in particular, the tradition established by Pröhle and other Turkologists who were writing for a different audience and a different purpose. (In fact, Pröhle himself did not invent the transcription system used in his glossary. He, too, was following an established tradition.) As one examines the whole series of Soviet publications, even dating back to the 1920's, when Karachay was written in the Roman alphabet, one finds exactly the same overdifferentiations that Pröhle used. It seems likely, then, that the Soviet scholars who devised writing systems for Karachay and the other Turkic languages were familiar with the tradition to which Pröhle belonged and borrowed the earlier transcription practices from that tradition. Incidentally, problems similar to the ones that we have seen with Karachay writing can be found in the spelling systems of many other Turkic languages of the Soviet Union, no doubt for the same reason.

The lesson to be learned from this case is that when one is writing a dictionary, one's choice of a transcription system for recording words should be determined by the intended audience of the dictionary. In a scholarly dictionary intended primarily for linguists and other researchers, a phonetic transcription may be (but is not necessarily) the appropriate choice. But in a practical dictionary intended for the speakers of the language, where the aim is to promote literacy, then a phonemic writing system, which matches the phonological intuitions of the speakers, is the best choice.

References

- Akbaev, Š. X. 1963. *Fonetika Dialektov Karačaevno-Balkarskogo Jazyka*. Čerkessk: Karačaevno-Čerkesskoe Knižnoe Izdatel'stvo.
- Aliev, U. B. 1972. *Sintaksis Karačaevno-Balkarskogo Jazyka*. Moscow: Nauka.
- Bayramkullany, A. and I. Orusbijlany. 1965. *Qaračay-Malkar Tilni Gramatikasy*, Čerkessk: Stavropol' Kitab Izdatel'stvo.
- Borovkov, A. 1935. 'Očerki Karačaevno-Balkarskoj Grammatiki' in G. P. Serdučenko (ed.). *Jazyki Severnogo Kavkaza i Dagestana*. Moscow-Leningrad: Gosudarstvennoe Social'no-Ekonomičeskoe Izdatel'stvo..
- Pröhle, Wilhelm, 1909a. 'Karatschajische Studien' in *Keleti Szemle*, vol. 10. Budapest.
- Pröhle, Wilhelm, 1909b. 'Karatschajisches Wörterverzeichnis' in *Keleti Szemle*, vol. 10. Budapest.
- Sujunčev, X. I. and Urusbiev, I. X. 1965. RUSSKO-KARAČAEVNO-BALKARSKIJ SLOVAR' ('Russian-Karachay dictionary'). Moscow: Sovetskaya Enciklopedia.
- Urusbiev, I. X. 1963. *Sprjaženie Glagola v Karačaevno-Balkarskom Jazyke*. Čerkessk: Karačaevno-Čerkesskoe Knižnoe Izdatel'stvo.
- Xabyčev, M. A. 1966. 'Karačaevno-Balkarskij Jazyk' in I. A. BASKAKOV (ed.). *Jazyki Narodov SSSR*, vol. 2, *Tjurkskie Jazyki*, Moscow: Nauka.