



Bitig Türkoloji Araştırmaları Dergisi, Bahar 2021 / 1: 37-50.

## South Siberian Material in Radloff's Dictionary\*

Kamil STACHOWSKI\*\*

**Abstract:** *South Siberian material makes up about a quarter of Wilhelm Radloff's Versuch eines Wörterbuches der Türk-Dialecte, making the dictionary by far the richest source for 19th century Siberian Turkic. The paper examines three aspects of this collection: its coverage, its phonetic accuracy, and the methodological choices made by Radloff, together with their implications for the usefulness of what is probably his most momentous work.*

**Keywords:** *Siberian Turkic, Versuch eines Wörterbuches der Türk-Dialecte.*

### **Radloff'un Sözlüğündeki Güney Sibirya Malzemesi**

**Öz:** *Güney Sibirya materyali, Wilhelm Radloff'un Versuch eines Wörterbuches der Türk-Dialecte'sinin yaklaşık dörtte birini oluşturmakta ve bu da sözlüğü 19. yüzyıl Sibirya Türkçesi için açık ara en zengin kaynak hâline getirmektedir. Bu makalede söz konusu koleksiyon üç yönden incelenmektedir: kapsam, fonetik doğruluğu ve Radloff tarafından yapılan metodolojik seçimler ile birlikte muhtemelen en önemli eserinin kullanılabilirliğine ilişkin çıkarımlar.*

**Anahtar Sözcükler:** *Sibirya Türkçesi, Versuch eines Wörterbuches der Türk-Dialecte.*

### **1. Introduction**

Wilhelm Radloff's *Versuch eines Wörterbuches der Türk-Dialecte*, along with Mahmud al-Kashgari's opus, is one of just two comprehensive, expressly comparative dictionaries of Turkic languages completed to date.<sup>1</sup> If the pace remains constant, we

---

\* I would like to express my gratitude to Mateusz Urban, PhD, Jagiellonian University, Cracow, Poland, for his tireless help with all things phonetic.

\*\* PhD, Jagiellonian University, Cracow, (Poland). E-mail: kamil.stachowski@gmail.com/ ORCID ID: 0000-0002-5909-035X.

1 KTLS, despite its name, is not a comparative dictionary. The task of a comparative dictionary is to prepare ground for etymological studies by connecting genetically related words into families, regardless of their contemporary meaning, e.g. Kzkh.dial. *kujy* 'a deep place in a river' = Tksh. *kuyu* 'well (hole), shaft' = Tuv. *kudu* 'lower, bottom' = Yak. *çotu* 'north'. KTLS's self-declared goal, on the other hand, is clearly that of a multilingual dictionary: "Türk dilinin zenginliğini ortaya koyacak ve Türk topluluklarının birbirlerini anlamalarına yardımcı olacak «Karşılaştırmalı Türk Lehçeleri Sözlüğü». [...] Hedefimiz, Türkiye'den Azerbaycan'a, Türkistan'a, İdil-Ural'a giden insanlar için bir kılavuz sözlük hazırlamaktır."

can expect that a new one will be published around the year 2840. This prediction will likely not come true because ÈSTJa is planned to be completed before this date, but regardless, *Versuch* will remain forever the single richest source of 19th century material (except Yakut and Chuvash). The present paper can be viewed as a practical guide to South Siberian material attested in it.

According to Radloff himself, a plan to compile a comprehensive dictionary of Turkic languages was already in his mind when he embarked on his scholarly career in 1859 (R I 1). It would be another 34 years before the first volume was eventually published. The dictionary began as a list of Altai and Teleut words, to which material from other languages was later added.<sup>2</sup> Young Radloff was helped and encouraged in his endeavour by Franz Anton Schiefner; it is to Schiefner's advice, informed surely by his interest in Tibetan, that the dictionary owes its alphabetical order (R I II–III, v). Radloff continues and, I might add, improves on the tradition of unique arrangements for comparative dictionaries of Turkic languages (cf. Stachowski, 2019: 228). Over the following years, Radloff seized every opportunity that presented itself to him, be it in the form of publication of new materials, or in the form of his twelve-year stay in Kazan, and used it to expand his dictionary (R I III–IV). It was probably only in late 1880s or early 1890s that, spurred on by his friends, Radloff decided to finally stop adding new material, and to publish his work (R I IV). Even when it was in print, however, he did in fact expand the Ottoman material one last time (R I v) and, although he could not have foreseen it while writing the foreword, the appearance in later volumes of languages that are not mentioned in the list of abbreviations (cf. footnotes in tab. 1) suggests that he did not resist further additions during the eighteen years that it took to publish the entirety of the dictionary.

One consequence of this methodology is that the dictionary is less than consistent in its coverage of various dialects; rather, it is simply the sum of knowledge available to Radloff. Indeed, he was perfectly aware of this, and deliberately used the word *Versuch* in the title to reflect this fact (R I IV–v). The story of Radloff's dictionary is nonetheless an inspiring tale which shows how thoroughly impressive a work can be completed almost in between other projects, simply through patience and persistence.

The present brief overview of South Siberian material in Radloff's dictionary is divided into three parts: sec. 2 discusses the breadth and depth of the coverage, sec. 3 its phonetic accuracy, and sec. 4 offers a handful of practical warnings. The whole is shortly summarized in sec. 5.

## 2. Coverage

Overall, Radloff's dictionary contains ca. 67,500 entries (von Gabain and Veenker, 1969–1972: VII). More than a quarter of this number, more than 18,000 entries, include South Siberian forms. It is not uncommon for one entry to feature multiple languages,

---

(KTLS xi). Accordingly, one will find in it families such as Az. *källä* 'skull' = Bshk. *baş höjägi* id. = Kirgh. *baş sōk* id. = Tksh. *kafatası* id. = Trkm. *kelle süŋki* id. These two approaches are not compatible.

2 Temir (1991: 97) recounts this story differently: "W. Radloff, önce yalnız topladığı metinlerdeki sözleri işine alan ve gezileri esnasında işine yarayan küçük, pratik bir «Lexikon» ile işe başlamıştı." Temir does not refer to any source here. Radloff, on the contrary, does not mention the practical aspect, and writes expressly "ein möglichst umfangreiches lexikalisches Sprachmaterial zusammen zu tragen." (R I 1).

so the total number of South Siberian words is considerably higher, nearly 31,000. Altogether, twenty-two South Siberian dialects are represented in the dictionary. The division of this group into languages and dialects remains a debated topic to this day, but it will be safe to say that with the exception of Fuyu Kyrgyz and Western Yugur, all the major branches are included: Chulym (*Chulym, Küärik*), Khakas (*Abakan, Beltir, Kacha, Kamasin, Koibal, Kysyl, Sagai*), Oïrot (*Altai, Chuya, Kumandin, Lebedin, Teleut, Telengit (= Tölös), Tuba*), Shor (*Kondoma, Matyr, Mrassu, Shor*), Tofa (= *Karagas*), and Tuvinian (= *Soyon*). (Regarding the names, cf. Baskakov, 1960: 230–238, and footnotes in tab. 1.) The extent of this representation, however, varies considerably. In sum, Oïrot dialects have as many as 16,447 forms, Khakas dialects 8065, Shor has 4590, Chulym 1285, Tuvinian only 345, and Tofa mere 11.

These numbers are only an approximation, however, as three types of uncertain entries have been omitted. The first type are 22 entries marked as “alle Dialecte” and one as “östl. Dial.” It is not clear whether this includes those dialects which are represented in the dictionary but are missing from the list of abbreviations (R I XVI–XVIII), e.g. *Matyr* or *Kamasin*; cf. sec. 4. The second type are 73 cases where a language is mentioned in the header of an entry but all the meanings are attributed to other languages, e.g. R II 758 *кырчыи* is said to be present in Kyrgyz, Lebedin, and Teleut, but it only has two meanings: one limited to Kyrgyz, the other to Teleut. The third type are at least 45 entries which are the inverse of type two, e.g. R I 1214 <sup>4</sup>õн is not attributed to Tuba but its meanings four and five are. (See also sec. 4.)

Dialect	Entries	Dialect	Entries
Teleut	8348	Kysyl	150
Shor	4491	Matyr <sup>3</sup>	96
Altai	4337	Tölös <sup>4</sup>	25
Sagai <sup>5</sup>	3921	Beltir <sup>6</sup>	22
Lebedin	2768	Karagas <sup>7</sup>	11
Koibal	2517	Chulym <sup>8</sup>	9
Kacha <sup>9</sup>	1451	Chuya <sup>10</sup>	5
Küärik	1276	Abakan	3

3 Abbreviations “Mad.” and “Mat.” are not explained, but some of the entries contain references to Verbickij (1884) where the same words are marked as “мат[ырский]”, a subdialect of Kondoma (Verbickij, 1884: iii).

4 Twenty entries are marked as “Tölös”, and further five as “Töl.”. The abbreviations are not explained, it is my guess that they refer to Tölös.

5 The number includes one entry where the abbreviation “Sag.” is followed by a question mark in brackets.

6 The abbreviation “Belt.” is not explained, it is my guess that it refers to Beltir.

7 The abbreviation for Karagas is “Karg.”, but the number includes one entry marked as “Karga”, as well as two marked with unexplained abbreviations “Karag.” and “Karakasch.” (one entry each).

8 One entry is marked as “Tscholym”, further seven as “Tschol.”, and one as “Tscholyschm” [sic]. The abbreviations are not explained, it is my guess that they refer to Chulym.

9 The abbreviation for Kacha is “Ktsch.”, but the number includes 2 entries marked with the unexplained abbreviation “Katsch.”, and one with the unexplained abbreviation “Katch.”.

10 Four entries are marked as “Tschuja”, and one as “Schuja”.

Kamil STACHOWSKI

Kumandin <sup>11</sup>	572	Kamasin <sup>12</sup>	2
Tuba <sup>13</sup>	391	Mrassu <sup>14</sup>	2
Soyon	345	Kondoma <sup>15</sup>	1
<b>Total</b>			<b>30,743</b>

Table 1. Number of entries attributed to different dialects.

With such a distribution, the semantic coverage cannot possibly be uniform. Since the preferred ‘round number’ in Siberia is nine (M. Stachowski, 2011), I prepared a sample of nine semantic groups, each with nine basic meanings in them:

- Active verbs: ‘to come’, ‘to run’, ‘to fly’, ‘to eat’, ‘to drink’, ‘to suck’, ‘to say’, ‘to give’, ‘to hunt’;
- Stative verbs: ‘to see’, ‘to hear’, ‘to think’, ‘to know’, ‘to sit’, ‘to stand’, ‘to sleep’, ‘to love’, ‘to fear’;
- Animate nature: ‘man’, ‘woman’, ‘dog’, ‘wolf’, ‘bear’, ‘fish’, ‘bird’, ‘tree’, ‘seed’;
- Body parts: ‘head’, ‘eye’, ‘nose’, ‘mouth’, ‘ear’, ‘leg’, ‘heart’, ‘bone’, ‘tail’;
- Family relations: ‘mother’, ‘father’, ‘daughter’, ‘son’, ‘younger sister’, ‘elder sister’, ‘younger brother’, ‘elder brother’, ‘husband’;
- Inanimate nature: ‘Sun’, ‘star’, ‘night’, ‘cloud’, ‘lake/sea’, ‘water’, ‘fire’, ‘stone’, ‘mountain’;
- Adjectives: ‘good’, ‘bad’, ‘big’, ‘small’, ‘cold’, ‘warm’, ‘long’, ‘new’, ‘red’;
- Pronouns: ‘I’, ‘thou’, ‘he/she/it’, ‘we’, ‘you’, ‘this’, ‘who?’, ‘what?’, ‘how?’;
- Numerals: ‘two’, ‘four’, ‘six’, ‘eight’, ‘ten’, ‘thirty’, ‘fifty’, ‘hundred’, ‘much/many’.

The sample was mostly informed by the Swadesh 200 list (Swadesh 1952: 456–457) with the expectation that all the most represented dialects up to Tuvinian (cf. tab. 1) will have the majority of meanings attested. It can be seen from tab. 2 that this is not the case.

We know from Radloff himself that his dictionary began as a list of Altai and Teleut words, which he then expanded over the years (cf. sec. 1). It can also be read between the lines of the foreword (R I I, III–V) that the expansions were not aimed at keeping the coverage of the dictionary consistent, but rather intended to make it as complete a collation of the available material as was possible. Tab. 2 supports this

---

11 The abbreviation for Kumandin is “Kumd.”, but the number includes 303 entries marked with the unexplained abbreviation “Kmd.”, as well as one where this abbreviation is followed by a reference to Verbickij (1884) and a question mark.

12 The abbreviation “Kamass.” is not explained, it is my guess that it refers to the language of the Kamasins, an originally Samoyedic tribe that was Turkified in late 19th / early 20th century (Tugarinov, 1926: 73, 83, 87; Wixman, [1984]: 91).

13 The number includes one entry marked with the unexplained abbreviation “Alat.”, as it also contains a reference to Verbickij (1884) where the word is marked as “алад[арское]” (= диалект Черневых татар).

14 Marked as “Mrass”.

15 Marked as “Kondoma”.

reading. Naturally, Radloff could not have possibly used the Swadesh list but the meanings I chose are all rather basic, and if he had followed a premeditated list, I suppose it would have included at least most of them. At any rate, I would have at least expected to see much less variation in tab. 2.

An alternative explanation for tab. 2 would be that the Swadesh list is not as useful a tool as it is often held to be. It is my conjecture that the majority of Radloff's sources were texts and wordlists based on texts. In other words, that the probability of a given meaning being attested correlates with its frequency in texts rather than with how basic a concept it may appear to be. This would support this alternative explanation, but of course it is also possible, and indeed seems to be quite likely, that both explanations are true at the same time.

One point about tab. 2, however, is not clear to me. In tab. 1, there are two pronounced drops: between Teleut and Shor, and a considerably smaller one between Küärik and Kumandin. The first one is not reflected in tab. 2 at all which can probably be explained by the fact that even the less represented Shor still has a relatively large number of words. On the other hand, the second drop is quite dramatic in tab. 2, and in addition followed by a slight rise in the number of attestations for Tuba and Tuvianian. I can think of no good explanation for this phenomenon.

	Teleut	Shor	Altai	Sagai	Lebedin	Koibal	Kacha	Kiärik	Kumandin	Tuba	Soyon	Kysyl
Active verbs	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Stative verbs	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Animate nature	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Body parts	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Family relations	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Inanimate nature	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Adjectives	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Pronouns	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿
Numerals	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿	⦿

Table 2. Number of meanings attested in different languages (each class contains nine meanings). Languages missing from the table have no meanings attested in them.

### 3. Accuracy

From the modern perspective, the transcription employed by Radloff would probably have to be classified as high-level phonetic, just one step below phonological ( $\approx$  L5 in K. Stachowski, 2011: 332), but Radloff himself makes no comment about this. It might be interesting to note that his stay in Kazan coincided with J.N. Baudoin de Courtenay's time at this university, and that this was the exact moment when the

latter introduced the distinction between phonetics and phonology in the modern sense (Jakobson, [1958]: 399). Radloff even attended de Courtenay's classes (Skarzyński, 2016: 29, fn. 21) and the two maintained a good relation (Blagova, 1979: 99–100), with conversations extending beyond linguistics (Baudoin de Courtenay, 1904), but apparently, the concept of “psychophonetics” was still too fresh to find practical application in the dictionary.

The set of characters used in the dictionary is a slightly unusual mixture of Cyrillic and Latin, which appears to have been largely dictated by the availability of types in Russian printing houses (for example, ⟨ö⟩ and ⟨ý⟩ sport different diacritics to denote the same modification). The meanings of individual characters are explained in some detail in the foreword, most often by comparing them to German, French and Russian sounds. The description is given in German and Russian simultaneously, but the two versions are not always identical. A translation to a contemporary system, such as the International Phonetic Alphabet (= IPA; [www.internationalphoneticassociation.org](http://www.internationalphoneticassociation.org)) or the Finno-Ugric transcription (=FUT; Sammallahti, 1998: 173–176; Sovijärvi and Peltola, 1977), is not a straightforward task and requires commentary that goes beyond the scope of the present paper. It can be found in K. Stachowski and Urban ([in preparation]).

Radloff took as it were a normative approach in his dictionary, in that each word is essentially only given a single phonetic shape per dialect. Exceptions to this rule are relatively rare.<sup>16</sup> This is in stark contrast to the great variation attested in Güner Dilek (2015). Radloff's decision must have been a conscious one, it cannot be blamed e.g. on technical limitations of his times, because in *Phonetik* he states clearly that some of the characters in his transcription represent multiple allophones simultaneously (Radloff, 1882: IX–XI; the transcription in *Phonetik* is quite compatible with that in *Versuch*, see K. Stachowski and Urban [in preparation]).

\*

As many as 26 different vowels can be found in South Siberian entries in Radloff's dictionary: ⟨a, á, ā, ä, ā, e, i, ī, ĭ, o, ó, ö, õ, õ, y, ý, ŷ, ŷ̄, ŷ̄̄, ы, Ы⟩. Of these, three are not explained in the foreword: ⟨á, ó, ý⟩. They appear in several borrowings from Russian. e.g. Tel. ⟨óсно⟩ ‘smallpox’ (< Russ. *óсна* id.; R I 1143), or Sag. ⟨капты́с⟩ ‘hat with a visor’ (< Russ. *капмы́з* id.; R II 201). Most likely, the acute denotes the place of stress. The phonological value of the remaining characters does not require explanation; for the phonetic value, as mentioned above, see K. Stachowski and Urban ([in preparation]).

A full, comparative analysis of Radloff's dictionary and newer sources far exceeds the scope of the present paper. It is, however, a happy coincidence that for Altai, the third best represented language, we have available a modern description that takes a lower-level, phonetic approach. Radloff and Güner Dilek (2015; =ATA) appear to generally agree on most of the major points, but even a brief comparison reveals discrepancies.

---

<sup>16</sup> Notably, the alternation ⟨j⟩- ~ ⟨j̣⟩- in Teleut (see sec. 4), and the alternation single ~ geminated voiceless stops in Altai and Teleut, e.g. R IV 607 *сы́ка* = R IV 612 *сы́кка* ‘to compel’.

In the area of short vowels, both R and ATA are approximately of the same mind about the values of /a, ä, i, o, ö, u, ü, y/, except ATA distinguishes eight additional allophones.<sup>17,18</sup> The only major discrepancy concerns narrow *e*, i.e. the opposition *ä* : *e*. Radloff distinguishes the two, although ⟨e⟩ only appears in five Altai words, versus 655 occurrences of ⟨ä⟩.<sup>19</sup> ATA likewise distinguishes ⟨è⟩, but in words in which Radloff notates ⟨ä⟩.<sup>20</sup> Meanwhile, Dyrenkova (1940: 14–17), Elcan (2019: 41), Kanaev (1931: 4–5), and Tybykova (1972) do not distinguish the two sounds at all.

Long vowels appear to be less controversial than short ones. Radloff and ATA generally agree on the values of /ā, ī, ō, õ, ū, ũ/, though not necessarily on their distribution.<sup>21</sup> Considering that Altai long vowels are mostly secondary, a side effect of contraction, these discrepancies likely only reflect the natural variation in what is primarily a non-literary language. Regarding /ā/, R I X remarks that it is pronounced very long in Altai, “almost like *aa*”; ATA transcribes it simply as IPA [ɑ:] (= FUT [ā]; ATA 18). Both Radloff and ATA only distinguish one *e*-type vowel for Altai, a long counterpart to wide *ä*, though they do not always do so in the same words.<sup>22</sup>

The most mysterious is the case of reduced vowels. They are not rare in Radloff's dictionary (120 occurrences of ⟨ī̄⟩, 28 of ⟨ÿ̄⟩, and 70 of ⟨ÿ̄̄⟩), and while phonetically

---

17 The term *allophone* is used here not only in the more common sense of sounds locked in a complementary distribution, but also to refer to sounds in free variation, one caused by personal preference of the speaker, resulting from dialectal diversification, etc.

18 This, and similar footnotes below have been created by extracting Altai examples from ATA's third chapter (ATA 93–156) and, where available, contrasting them with the same words in Radloff's dictionary.

⟨á⟩: ATA 103 ⟨áyda⟩ : R I 49 ⟨а́йда⟩ ‘to drive’.

⟨á⟩: ATA 102 ⟨gerák⟩ : R II 1086 ⟨kápäk⟩ ‘necessary’. ⟨è⟩: ATA 106 ⟨nème⟩ : R III 690 ⟨nämä⟩ ‘what?’, plus three more examples to this effect.

⟨é⟩: ATA 107 ⟨telgên⟩ ‘beginner’ missing from R.

⟨í⟩: ATA 109 ⟨gíyis⟩ ‘felt’ missing from R.

⟨ï̄⟩: ATA 110 ⟨doltıra⟩ : R III 1204 ⟨толтыра⟩ ‘completely’.

⟨ī̄⟩: ATA 112 ⟨ište⟩ : R I 1561 ⟨иштä⟩ ‘to work’, plus two more examples to this effect.

⟨ó⟩: ATA 113 ⟨góc⟩ : R II 1287 ⟨коч⟩ ‘to live as a nomad’, plus two more examples to this effect.

19 R II 1119 ⟨kälışh⟩ ‘to fit’, R II 1134 ⟨kedäpä⟩ ‘away, aside’, R III 1973 ⟨çertäk⟩ ‘punch on the nose’, R III 1985 ⟨chedän⟩ ‘fence, hedge’, and R IV 1230 ⟨peripäk⟩ ‘a little closer’. The last one might be a misspelling, as R deconstructs it to *näpi* + *päk*.

20 ATA 106, 388 ⟨cér⟩: R III 334 ⟨jäp⟩ ‘place, land’; ATA 106 ⟨dè⟩ : R III 1008 ⟨tä⟩ ‘to say’; ATA 107 ⟨mèn⟩ : R IV 2085 ⟨mäñ⟩ ‘I’; ATA 106 ⟨nème⟩ : R III 690 ⟨nämä⟩ ‘what?’. However, cf. also ATA 105, 387 ⟨cer⟩, ATA 407 ⟨de⟩, ATA 475 ⟨neme⟩, and ATA 468 ⟨men⟩.

21 /ā/: ATA 104 ⟨cāñh⟩ : R III 121 ⟨jaryk⟩ ‘bright’, plus one more example to this effect.

/ī/: ATA 97 ⟨cīt⟩ : R III 523 ⟨jīt⟩ ‘youth’.

/õ/: ATA 98 ⟨ōdm⟩ : R I 1126 ⟨одун⟩ ‘firewood’, plus one more example to this effect.

/ȭ/: ATA 99 ⟨bõrū⟩ : R IV 1302 ⟨põpý⟩ ‘wolf’.

/ū/: ATA 99 ⟨ül⟩ : R I 1674 ⟨ýл⟩ ‘son’, plus one more example to this effect.

/ũ̄/: ATA 100 ⟨sũrt⟩ : R IV 825 ⟨cýрт⟩ ‘to cause to drive’.

22 ATA 96, 388 ⟨cēñ⟩ : R III 328 ⟨jāñ⟩ ‘nephew’; ATA 96 ⟨ēk⟩ : R I 676 ⟨āk⟩ ‘chin’ versus ATA 96, 388 ⟨cēr⟩ : R III 334 ⟨jāp⟩ ‘place, land’, but cf. also ATA 105, 387 ⟨cer⟩.

corresponding allophones can be identified in ATA, they do not appear to be used in the same words. It seems that the latter consistently gives plain short vowels in place of Radloff's reduced ones.<sup>23</sup> Dyrenkova (1940: 14–17), Kanaev (1931: 4–5), and Tybykova (1972) also do not distinguish such a category of vowels.

\*

The set of consonants is larger than that of vowels, counting 32 items: <б, ц, ч, д, ф, г, ғ, х, Һ, Җ, Ҙ, җ, ҝ, к, л, л, м, н, Һ, Һ, п, р, с, ш, т, в, w, з, 3, ж, ц>. In addition to these, R I XV announces the grapheme <б̄>, a sound intermediate between [b] and [m], which is supposed to only appear in the word <еб̄> in “Abakan dialects”. There is only one word that almost fits this description, but it is spelt <еб>, and attributed to Koibal and Sagai (R I 925).

In Altai words, 25 consonants are attested, and their phonetic value is more often than not the same or very similar as in ATA. This is the case with /b, č, d, d́, γ, h, j, m, n, ɲ, p, r, s, š, t, t́, z, ž, ž́/.<sup>24</sup> There are differences in the distribution, perhaps most notably in the distribution of initial stops. With the exception of /d́/ (see below), Radloff's dictionary only has voiceless variants in the anlaut; ATA attests both, voiceless and voiced ones, but the latter are considerably more numerous.<sup>25</sup> The discrepancy regarding *b-* may perhaps be put down to dialectal variation within Altai (voiced in Onguday, voiceless elsewhere; Dyrenkova, 1940: 34).

Moreover, ATA notates four allophones that are missing from Altai words in Radloff.<sup>26</sup> The treatment of the *k, g, l* triad is inconsistent in both sources. Radloff only distinguishes front and back variants of *k* and *l* while ATA only those of *k* and *g*.<sup>27</sup> The two sources also disagree about the phonetic values, Radloff failing to notice the

23 <й̄>: ATA 386 <cedinçi> : R III 366 <јәттинчи> ‘seventh’; ATA 428 <giyim> : R II 1344 <кijim> ‘clothing’; ATA 445 <çiydir> : R II 1378 <кідір> ‘to dress (trans.)’.

<й̄>: ATA 401 <çuluḥ> : R III 2176 <чулұк> ‘sock’; ATA 417 <durğus> : R III 1458 <турьҗус> ‘to put down’; ATA 516 <uçun> : R I 1725 <учұн> ‘for’.

<й̄>: ATA 393 <cüsün> : R III 620 <јүзүһн> ‘various’; ATA 490 <öttür> : R I 1255 <өлтүр> ‘to kill’; ATA 518 <üçüncü> : R I 1874 <үчүһнчи> ‘third’.

24 The case of /ç/ is not in fact entirely clear, but it is probably due to a mistake in ATA. According to ATA 19, <ç> represents the ‘[n]ormal ç ünsüzü’ which I understand to mean [tʃ]; this agrees with the description on p. 138. However, in the same table on p. 19, <ç> is simultaneously said to correspond to IPA [tʃ], and to be voiced. I believe the latter to be a mistake.

25 [b-]: ATA 119 <biyik> : R IV 1325 <пik> ‘high’, plus five more examples to this effect.

[d-]: ATA 126 <dalḥan> ~ <dalqan> : R III 889 <талқан> ‘roasted barley’, plus seven more examples to this effect.

[g-]: ATA 148 <ḡel> : R II 1109 <кäl> ‘to come’, plus one more example to this effect.

26 <ф>: ATA 125 <ḡıfçaḥ> ‘Kipchak’ missing from R.

<н̄>: ATA 145 <né> : R III 666 <nä> ‘what?’.

<v>: ATA 125 <var> : R IV 1145 <пар> ‘to reach, to go’, plus one more example to this effect.

<w>: ATA 121 <dawışdır> : R III 975 <табыштыр> ‘to deliver’; ATA 121 <wala> : R IV 1161 <пала> ‘child’.

27 Technically, ATA distinguishes between <ɫ> and <ɫ́> but both point to the same IPA [l] (= FUT [l]; ATA 20), and the former is used in both front and back words, while the latter is said to primarily appear in borrowings from Russian, and only “daḡmık bir şekilde” in native words (ATA 142). In addition, ATA distinguishes an allophone <ɫ̣> = IPA [ḷ] (?).



palatalization of front *k*, *g*, and ATA the darkness of back *l*. Moreover, Radloff notes that the exact place of articulation of back *k* varies between dialects, but it is always “much farther back than in German” (R I XIV); ATA records simply IPA, FUT [k]. The two agree about the strong aspiration of back *k*, only in Radloff's dictionary it is not clear that the remark includes Altai (“some Eastern dialects”; R I XIV), and in ATA the aspirated variant is treated as a separate allophone – together with five others that ATA distinguishes and Radloff does not.<sup>28</sup> It is a little surprising, in light of the comparisons made above, that Radloff's dictionary distinguishes three more sounds for Altai, which appear to be missing from ATA. One is ⟨ɣ⟩ and has no direct phonetic equivalent in ATA, but seems to correlate with ATA's ⟨g⟩ and ⟨ġ⟩.<sup>29</sup> The other two are ⟨j⟩ and ⟨y⟩; they are phonetically very similar to ATA ⟨d̥⟩ and ⟨t̥⟩ (IPA [d̥, t̥], FUT [d̥, t̥]), but they do not occur in the same words.<sup>30</sup>

#### 4. Caveats

In a work of such magnitude as Radloff's dictionary, mistakes are unavoidable. Some practices, however, are used in it quite frequently, and others almost regularly, which suggests that they were intentional for at least some time. They are not without consequence for the usefulness of the dictionary, and as such they need to be briefly discussed here.

It was mentioned in sec. 1 that the alphabetical order was not Radloff's idea, and it seems that he was himself not entirely sure how to handle it (R I III, v–vi). Misplaced entries are not very rare, but far more common is the unification of various letters. I suspect it was the treatment of unlauded letters in German that inspired Radloff to sort short, long, and reduced vowels all together (e.g. the very first entry in the dictionary, under *a*, is <sup>1</sup>ā), but the actual problem is that he was not consistent in which letters he unified. For example, according to R I IX, the order should be *k* – *κ* – *x* – *ɣ* – *z* – *z̄* – *ɣ*; the actual order of sections is *k* – *κ*, *ḱ* – *ɣ* – *z*, *z̄* – *x* (no words begin with *z̄*- or *ɣ*-), but in further volumes some of those are sometimes sorted together anyway, e.g. in R III 1245 <sup>1</sup>mözyl, <sup>2</sup>möhyl.

Mistakes in the alphabetical order result sometimes in homophonic entries not being numbered when they become separated by another entry or entries (e.g. *nocmo* in R IV 1291 and 1292). In fact, numbers are sometimes missing even when the order is correct and the two homophones are directly one after the other (e.g. in R III 846 there

28 ⟨ġ⟩: ATA 148 ⟨ġiyim⟩ : R II 1344 ⟨κijim⟩ ‘clothing’, plus two more examples to this effect. Note that ⟨ġ⟩ and ⟨g⟩ point to the same IPA [j] (= FUT [ġ]; ATA 19), but are not in fact identical, cf. ATA 148 and 149.

⟨g⟩: ATA 147 ⟨gós⟩ : R II 1291 ⟨κös⟩ ‘eye’, plus seven more examples to this effect.

⟨h⟩: ATA 153 ⟨halan⟩ : R II 230 ⟨калан⟩ ‘tax’, plus one more example to this effect.

⟨k̄⟩: ATA 146 ⟨kúc⟩ : R II 1489 ⟨куч⟩ ‘strength, power’, plus four more examples to this effect, although cf. also ATA 146 ⟨ekel⟩ : R I 682 ⟨äkkäl⟩ ‘to bring’.

⟨q⟩: ATA 152 ⟨alqıš⟩ : R I 391 ⟨алқыш⟩ ‘blessing’, plus two more examples to this effect.

29 ⟨ɣ⟩: ATA 393 ⟨cügür⟩ : R III 595 ⟨j'ýb'ýp'⟩ ‘to run’; ATA 398 ⟨cegen⟩ ~ ⟨ceġen⟩ : R III 1957 ⟨чәһәән⟩ ‘kumis’; ATA 461 ⟨körgüz⟩ : R II 1263 ⟨көрһүс⟩ ‘to show’.

30 Both ⟨j⟩ and ⟨y⟩ are rare, the former being attested in ten Altai words, the latter in eight. I was only able to correlate two stems between R and ATA:

⟨j⟩: ATA 418 ⟨deke⟩ : R III 1016 ⟨täkä⟩ ‘goat’.

⟨y⟩: ATA 474 ⟨nacı⟩ : R III 644 ⟨наҗы⟩ ‘friend’.

are as many as five unnumbered *тары́*'s). This is an inconvenience, but I do not believe that it has great negative impact on the usefulness of the dictionary as a whole. A more serious problem is when an entry contains nothing more than the phonetic shape, the list of dialects, and a reference to another entry – an unnumbered reference when the referred entry has multiple homophones; such is e.g. is the case with R I 1158 *обак* which refers to just *омак*, but there are in fact two different *омак*'s in R I 1166. This is not a rare issue, it affects about a quarter of all references.

One example that combines all of these problems at the same time, is R IV 1408 *пуду́р҃л*. The relevant entries are as follows:

<sup>1</sup>пуду́р [Tel.]

неровность — die Unebenheit.

<sup>2</sup>пуду́р (v) [[...] Tel. [...], von пұт+р]

совершать, готовить, творить — vollenden, fertig machen, verfertigen, schaffen [...]

[...]

пуду́р҃л (v) [Tel., Alt., von пұдұр+л]

быть оконченнымъ, исполненнымъ — geendigt, ausgeführt werden.

пуду́рлӓ (v) [Tel., von <sup>1</sup>пуду́р+лӓ]

спотыкаться, запинаться, зацѣпляться — hängen bleiben, anstossen, stolpern.

пуду́р҃л (v) [Tel.] = пуду́рлӓ

As can be seen, the mystery of the two *пуду́р҃л*'s unravels with relative ease, but not all entries are so cooperative.

Continuing the theme of references, it is not very uncommon that they point to non-existent entries. The mistake is usually in the quantity of the vowel, e.g. R IV 1303 <sup>1</sup>*nöly̆c* refers to *nöly̆u*, and <sup>2</sup>*nöly̆c* to <sup>2</sup>*nöly̆u*, but an unnumbered *nöly̆u* does not exist, and neither does <sup>2</sup>*nöly̆u*. Phonetically, the closest entries are <sup>1</sup>*nöly̆u* and <sup>2</sup>*nöly̆u* in the same column, but since the references do not contain meanings, and the two *nöly̆u*'s do not contain examples of usage, it is virtually impossible to resolve the reference to the unnumbered *nöly̆u*, and to tell whether the *-ü-* is reduced or not – or, indeed, whether the resolution to *nöly̆u* is in fact correct.

It should be also noted that the use of references appears to be generally somewhat random. When there is a family of related words, ideally, I would expect either all entries to point to one selected entry where references to all the other forms can be found, or alternately, all entries to point to all the other entries. This is rarely the case in Radloff's dictionary; an example, using the word for 'swallow', is in fig. 1 (including all languages, not only South Siberian ones).

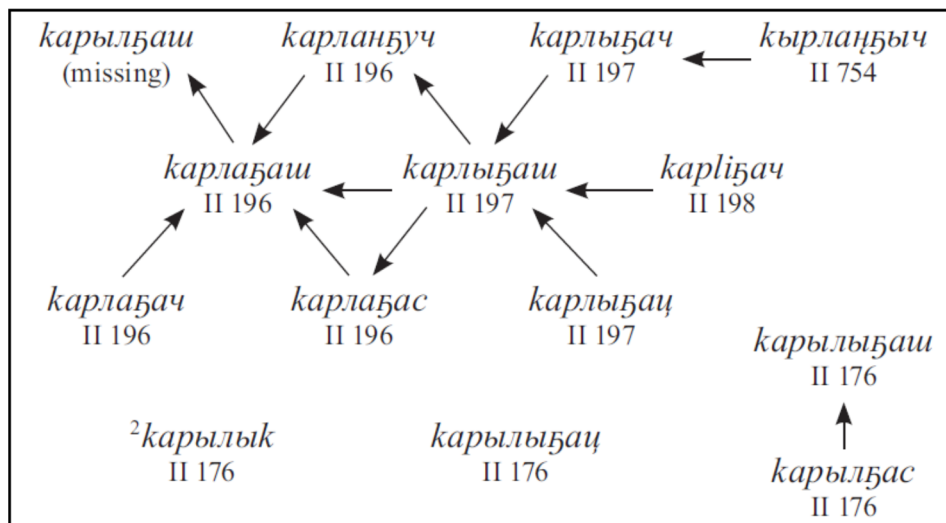


Figure 1. References between entries with related words for ‘swallow’.

There is, however, one moment where seemingly faulty references are not just perfectly correct but can even prove very helpful. I cannot stress enough how easy the distinction between ⟨j⟩ and ⟨j̣⟩ is to overlook and the two letters to confuse, and since they are sorted together, it helps to draw attention to this minuscule difference when one encounters such pairs as e.g.

<sup>1</sup>jaikындан (v) [Tel. Alt., von jaikын+лан]

разливаться (о рѣжахъ) — überschwemmt sein [...]

<sup>2</sup>jaikындан (v) [Tel.] = jaikындан.

[...]

<sup>1</sup>jaikkындан (v) [Tel. Alt.] = jaikындан.

<sup>2</sup>jaikkындан (v) [Tel. Alt.] = jaikындан.

Still, some references are mistakes, e.g. R III 1254 “töröt (v) [Tel. Alt.] = töröt” (there is no other *töröt*, or at least not anywhere near, and not anywhere within South Siberian entries).

Another relatively common failing, mentioned already in sec. 2, concerns attribution to dialects. Probably the most frequent variant is when a language is missing from the list at the beginning of the entry, but appears in one of the meanings. For example, R II 82 *качыр* is only attributed to Tatar and Teleut, but its second meaning is attested in Beltir, Kacha, Koibal, Tatar, and Sagai. The inverse is rarer but also possible, e.g. R II 204 *карчы* is assigned to Lebedin, Koibal, Sagai, and Teleut, but then its first meaning is only attested for Sagai, Teleut, and “Miss.” (abbreviation unexplained), and its second and last meaning only for Sagai and Koibal. It is very rare for an entry to not be attributed to any language at all (e.g. R IV 254 *саклаңи*). Sometimes, the problem only becomes apparent when several entries are compared.

For example, in R IV 289, the word *cajak* is attributed to Kumandin and Teleut, but not to Shor. A derivative from it, *cajakma* is assigned to Shor, but not to Kumandin. However, *cajakmau*, a derivative from *cajakma*, is apparently present in Kumandin. This last example becomes more understandable when one assumes that Radloff's method was, as suggested in sec. 1, an ambition to include all the materials available to him, with limited regard for consistency.

There is also the problem of unresolved abbreviations. Radloff appears to have kept expanding his dictionary while it was already in print (cf. sec. 1 and 2) but he did not update the list of abbreviations. This means not only that the reader is forced to simply guess the meanings of some of the abbreviations, but also that when “alle Dialecte” or “östl. Dial.” appear in the list of languages, it is not clear which dialects are in fact included. South Siberian languages appear to be relatively lightly affected by these two problems, especially in comparison to Karakhanid dialects (cf. “Chami.”, “Chot.”, “Kutsch.”, “Log.”, “Turf.”, and others).

Lastly, etymologies are given in a very inconsistent way. On the one hand, nearly a third of South Siberian entries are described as derivatives (as e.g. *cajakmau* in the example above); on the other, entire families of closely related words are often left with no explanation at all (e.g. R III 369f *ʼjäädäk* ‘kind of rope’ – *jäädäklä* ‘to lead by a rope’ – *jäädäklän* ‘to be lead by a rope’). Borrowings from Russian tend to be marked as such, Mongolian loanwords are also often indicated, but in general, more demanding etymologies are exceptionally rare. Also, and this is something that cannot possibly be held against Radloff, but is nevertheless a problem that needs to be borne in mind, some of Radloff's etymologies have by now become obsolete (for examples, see Ölmez, 1997: 372–374).

## 5. Summary

Radloff's *Versuch eines Wörterbuches der Türk-Dialecte* is the only complete, modern comparative dictionary of Turkic languages. It started as an Altai and Teleut wordlist in 1859, and was continuously expanded over the following decades until the publication of the last volume in 1911. Radloff's primary goal, it seems, was to amass in it the entirety of Turkic material available. The result is one of the largest Turkic dictionaries today, with ca. 67,500 entries in it (Gabain and Veenker, 1969–1972: VII), which translates to probably more than 110,000 words.<sup>31</sup> However, it is at the same time a dictionary with a very uneven coverage. As many as 22 South Siberian languages are represented in it, but the majority have fewer than a thousand words. It is also for each language a separate, different set of words; cooccurrences of the same stems or meanings across multiple dialects are, it appears, coincidental.

The transcription used by Radloff is high-level phonetic, nearly phonological. The dictionary distinguishes few allophones, one result being that alternative pronunciations within a single dialect generally tend to go unrecorded. Surprisingly, however, it recognizes several allophones – phonemes? – that a modern description does not. Further, detailed research is necessary to reconcile the two sources.

---

31 There are 18,241 entries with South Siberian words in them, and the total number of South Siberian words is 30,743 (cf. sec. 2). Assuming that the ratio for all the other languages is the same, the total number of words should be ca. 113,763.

Radloff's decision to focus on fullness rather than consistency means that only a part of his dictionary can find immediate application in comparative studies, the part that is the intersection of the different sets of words attested for various languages. Indeed, unusual methodological choices do not end there, but what is perhaps most damaging is Radloff's apparent lack of determination to adhere to those choices, once he had made them, throughout the dictionary.

Naturally, such criticisms are much easier to voice when one has the entire work ready and printed on one's desk, than they are to prevent over decades of labour. Despite its quirks and pitfalls, *Versuch* will remain forever a hugely important source, and not merely because it is by far the largest from the 19th century, but simply because it is, all in all, a good and vast dictionary.

### Abbreviations and references

**Az.** = Azerbaijani; **Bshk.** = Bashkir; **dial.** = dialectal; **Kirgh.** = Kirghiz; **Kzkh.** = Kazakh; **Russ.** = Russian; **Sag.** = Sagai; **Tel.** = Teleut; **Trkm.** = Turkmen; **Tuv.** = Tuvinian; **Yak.** = Yakut

ATA = Güner Dilek 2015.

Baudoin de Courtenay J. N. (1904). Z logiki tłumów. *Krytyka*, 6(1). 371–372.

Baskakov 1960 = Баскаков, Н. А. (1960). *Тюркские языки*. Москва: Издательство восточной литературы.

Baskakov 1966 = Баскаков, Н. А. (1966). *Диалект черневых татар (Туба-кижи). Грамматический очерк и словарь*. Москва: Наука.

Vlagoва 1979 = Благова, Г. Ф. (1979). Памяти И.А. Бодуэна де Куртене. *Советская тюркология*, 1979(6), 99–101.

Dyrenkova 1940 = Дыренкова, Н. П. (1940). *Грамматика ойротского языка*. Москва, Ленинград: Издательство Академии Наук СССР.

Elcan, A. (2019). *Altay Türkçesi. Karşılaştırmalı Ses ve Şekil Bilgisi*. New York, Ankara: Gece Kitaplığı.

Ercilasun, A. B. et al. (1991–1992). *Karşılaştırmalı Türk Lehçeleri Sözlüğü*. Ankara: Kültür Bakanlığı.

ÈSTJa = Sevortjan et al. 1974–.

Güner Dilek, F. (2015). *Güney Sibiryaya Altay Türkçesi Ağızları (Altay Kişi, Telengit Kişi, Çalkandı Kişi, Kumandı Kişi, Tuba Kişi). Karşılaştırmalı Ses Bilgisi*. Ankara: Türk Dil Kurumu.

Jakobson, R. [1958]. The Kazan' school of Polish linguistics and its place in the international development of phonology. *Roman Jakobson. Selected Writings II: Word and Language in (394–428. pp.)*. The Hague, Paris: Mouton.

Канаев <sup>3</sup>1931 = Канаев (Тоцаков), М. С. (1931). *Грамматика ойротского (алтайского) языка*. Ulalu: Ojrotizdat.

KTLS = Ercilasun et al. 1991–1992.

Ölmez, M. (1997). Radloff Sözlüğünün Yeni Bir Yayımı Nasıl Olmalıdır?. K. İmer and N. E. Uzun (eds.), *VIII Uluslararası Türk Dilbilimi Kurultayı Bildirileri. 7–9 Ağustos 1996*. Ankara: Ankara Üniversitesi.

R = Радловъ, В.В. / Radloff, W. (1893–1911). *Опытъ словаря тюркскихъ нарѣчій. Versuch eines Wörterbuches der Türk-Dialecte*. Санктпетербургъ: Императорская академія наукъ.

Radloff, W. (1882). *Phonetik der nördlichen Türksprachen*. Leipzig: T.O. Weigel's Verlag.

Sammallahti, P. (1998). *The Saami Languages. An Introduction*. Kárášjohka: Davvi Girji.

Sevortjan et al. 1974– = Севортян, Э. В. et al. (1974–). *Этимологический словарь тюркских языков*. Москва: Наука.

Skarżyński, M. (Ed.). (2016). *Materiały do dziejów polskiego językoznawstwa II: Jan Baudoin*

Kamil STACHOWSKI

- de Courtenay. Teksty mniej znane.* Kraków: Księgarnia Akademicka.
- Sovijärvi, A. and Peltola, R. (eds.). <sup>6</sup>(1977). *Suomalais-ugrilainen tarkekirjoitus.* Helsinki: [Helsingin yliopisto].
- Stachowski, K. (2011). Remarks on the usefulness of different types of transcription with a particular regard to Turkic comparative studies. *Suomalais-Ugrilaisen Seuran aikakauskirja / Journal de la Société Finno-Ougrienne*, 93, 303–338.
- Stachowski, K. and Urban, M. [in preparation]. Radloff's transcription.
- Stachowski, M. (2011). Siberian languages in contact, 2: The nonal counting system in Siberia and Russian *devjanosto* '90'. *Incontri Linguistici*, 34, 109–125.
- Stachowski, M. (2019). The Turkic Languages and Persian to c. 1700. J. Considine (ed.), *The Cambridge World History of Lexicography* in (223–246. pp.). Cambridge: Cambridge University Press.
- Swadesh, M. (1952). Lexico-Statistic Dating of Prehistoric Ethnic Contacts: With Special Reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society*, 96(4), 452–463.
- Temir, A. (1991). *Türkoloji Tarihinde Wilhelm Radloff devri. Hayatı – İlmî Kişiliği – Eserleri.* Ankara: Atatürk Kültür, Dil Ve Tarih Yüksek Kurumu.
- Tugarinov 1926 = Тугаринов, А. Я. (1926). Последние Калмажи. *Северная Азия*, 1926/1, 73–88.
- Тубукова 1972 = Тыбыкова, А. (1972). Об усовершенствовании и унификации алфавита алтайского языка. Н. А. Баскаков (ed.), *Вопросы совершенствования алфавитов тюркских языков СССР* in (41–48. pp.). Москва: Наука.
- Verbickij 1884 = Вербицкий, П. В. (1884). *Словарь алтайского и аладагского наречий тюркского языка.* Казань: Православное миссионерское общество.
- von Gabain, A. and Veenker, W. (1969–1972). "Radloff". *Index der deutschen Bedeutungen.* Wiesbaden: Harrassowitz Verlag.
- Wixman, R. [1984]. *The Peoples of the USSR. An Ethnographic Handbook.* London, New York 2015: Routledge.