



Accuracy Analysis of Mansi Transcription in  
Various Sources of the Beginning of the 20th  
Century.

---

Yulia Normanskaya and Natalia Koshelyuk

EasyChair preprints are intended for rapid  
dissemination of research results and are  
integrated with the rest of EasyChair.

September 15, 2022

**Accuracy analysis of Mansi transcription  
in various sources of the beginning of the 20th century.**

As early as the beginning of the 20th century, the Ob Ugric speakers of the Khanty and Mansi languages occupied a vast territory from the upper Pechora in the north of the Ural Mountains to the Tomsk region of the Yugan, Vasyugan and Vakh rivers (about three thousand kilometres from the northwestern to the southeastern regions of their settlement). Of course, the languages that spread over such a vast territory were characterised by significant dialectal fragmentation. The Khanty and Mansi languages are divided into four dialect groups each, between the speakers of which there is no mutual understanding. At the beginning of the 20th century, in each dialect group there were several more dialects, which also differed significantly from each other in morphological and phonetic systems. It turned out that even one dialect, for example, Sosva Mansi, common in the settlements of the village of Sosva and the village of Lombovozh, the distance between which is about a thousand kilometres, has a significant variability in dialects. As shown by recent studies conducted by E. V. Kovaleva in the phonetic program Praat, in the dialect of the village of Sosva there are long and short vowels, as in the literary Mansi, and in the village of Lombovozh, which is located much to the south, in the Sverdlovsk region, judging by analysis of archival data collected by Novosibirsk scientists in the 70s of the XX century, there are only neutral vowels. Thus, the Khanty and Mansi dialects at the beginning of the 20th century could be described as eight languages with several dozen dialects that differ significantly from each other.

Unfortunately, at present the situation is catastrophically changing. Some of the dialect groups have already disappeared: the last speakers of the southern and western dialects of the Mansi language, the southern dialects of the Khanty language, died already in the middle of the 20th century. Now only one speaker of the East Mansi dialect is alive (see below for more details). Some dialects of the Khanty language: Nizyam (an intermediate group between the northern and southern Khanty dialects), Salym (an intermediate group between the western and eastern Khanty dialects) were considered to have disappeared, but field expeditions carried out within the framework of our projects by S. V. Onina and M. K. Amelina (for more details see below), made it possible to find the last few speakers of these dialects.

According to the 2010 census, about 9.5 thousand people speak the Khanty language. In fact, we are actually talking only about the speakers of the northwestern dialects, because, as mentioned above, the southwestern dialects have already disappeared, and the eastern dialects are spoken only by individual speakers living in villages along the tributaries of the Ob, whose average age is about 70–80 years. Eastern Khanty dialects are not used in everyday life.

According to the 2010 census, about 4.5 thousand people speak the Mansi language. These are almost exclusively speakers of the northern Sosva and Sygvin dialects, since, as mentioned above, there are no longer speakers of the western and southern dialects, and we know only one speaker of the eastern dialect.

Therefore, for the western and southern dialects of the Mansi language, up to the present time, the materials of only two scientists, collected in the 19th century, have entered into

scientific circulation, and to some extent contradictory like [MK 1986], created based on the materials of B. Munkacsy and [Kannisto 2013] - based on the data of A. Kannisto. Are there big differences between these dictionaries, whose transcription can be relied upon? Let us look at one example that reflects the general picture quite well - there are the two words 'one' and 'autumn', which, according to V. Steinitz [Steinitz 1955], L. Honti [Honti 1982] and M. A. Zhivlov [Zhivlov 2006] have the same parent language vowel \*ü in the first syllable.

Table 1  
Comparing transcription done by A. Kannisto and B. Munkacsy

[MK 1986]	[Kannisto 2013]
N <i>akw</i> (~ <i>au</i> ~ <i>ak</i> ), LM <i>ākw</i> (~ <i>āu</i> ~ <i>āk</i> ), LU <i>āk</i> (~ <i>āk</i> ), P <i>əkw</i> ~ <i>əu</i> , K <i>ākw</i> ~ <i>ākwé</i> , T <i>uχ</i> - ( <i>uk</i> -) 'one' [MK 1986: 28]	T <i>üχ</i> , KM <i>āk</i> , P <i>ək</i> , VNK <i>äk</i> , LO <i>ākβā</i> , So <i>ḡGβā</i> 'one' [Kannisto 2013: 15]
N <i>takwès</i> , LM <i>tākwès</i> , LU <i>tākwès</i> , <i>tāχwès</i> , K <i>tāχwès</i> , T <i>tukès</i> (~ <i>tūkès</i> ) 'autumn'[MK 1986: 619]	KM, VNK <i>tākφs</i> , P <i>təkəs</i> , <i>tεχs</i> , So <i>tḡGφs</i> [Kannisto 2013: 363]

As can be seen, the recording of the first syllable vowel in dialects differs not only in two dictionaries for the same word, but also in different words of the same author. In particular, [MK 1986] has different vowels in Pelym (P *əkw* ~ *əu* 'one' and P *tākwès* 'autumn') and in Tavda (T *uχ*- (*uk*-) 'one' and T *tukès* (~ *tūkès*) 'autumn') dialects. According to [Kannisto 2013], there are also differences in these words, though not only in Tavda and Pelym, but in almost all dialects, with the exception of Sosva.

In 2015, while working at the National Library of Finland in Helsinki, a dictionary of the Upper Pelym dialect (Western Mansi) was found [Slotvsov 1905]. This dictionary contains 424 Western Mansi lexemes, for some separate inflectional forms are given. This dictionary was typed by us and is also currently available in electronic form on the site <http://lingvodoc.ispras.ru/>.

In this work, we plan to conduct a comparative analysis of these vowel graphemes of the first syllable in the Pelym dialect of the Mansi language, presented in dictionaries [Slotvsov 1905; MK: 1986; Kannisto 2013].

In order to verify the results obtained, we considered it expedient not only to compare dictionaries with each other lexemically, but also to involve modern field audio material processed in the Praat phonetic program to clarify the transcription.

In March 2013, M.K. Amelina had conducted an expedition to the village of Shugur in the Kondinsky district of the Khanty-Mansi Autonomous Okrug in order to collect material on the Yukondin dialect of the Mansi language, which, according to B. Munkachi's classification, is considered eastern. During the expedition, only two speakers of this dialect were found who could speak it: Selivanova (nee Nokhova) Elizaveta (Anna)<sup>2</sup> Ivanovna (b. 1923, at the time of recording - 89 years old) and Shivtorov Maxim Semenovich (1939 p., at the time of recording - 74 years old); the rest of the Mansi in the village of Shugur remember

only a few words. From these two native speakers, complete dictionaries of the original lexicon with paradigms and contexts of their use, as well as small texts, were recorded. Five days after the departure of M. K. Amelina from the village of Shugur, on April 2, 2013, Selivanova Elizaveta Ivanovna died.

It is important to note that at present, the analysis of data collected from two speakers of the East Mansi (Yukondin) dialect indicates their high significance for linguistics: it turned out that, compared with the literary North Mansi, the East Mansi dialect has archaic features - three degrees of vowel length and three accent paradigms of the verb. The collected material confirms that these features, which are also found in the Ob (Northern Mansi) dialect, should be reconstructed for the Proto-Mansi language as well. The presence of records from two carriers indicates that this or that feature is not a feature of the idiolect of one carrier, but was characteristic of the idiom as a whole. This example shows that every year of delay can be fatal for the preservation of the material of languages and dialects that are on the verge of extinction.

Previously, the phonetic differences between the Yukondin dialect and the Mansi literary language were described in [Kuzakova 1963; Saynakhova 2012], but the accent system in this dialect has not been studied. As a result of a complete analysis of the collected audio material, cut into separate word forms, it was found in the Praat program that three accent paradigms of the verb are distinguished in the Yukondin dialect of the Mansi language, see [Normanskaya 2015a].

So, in 2013, I. A. Stenin had conducted an expedition to two villages in the Oktyabrsky district of the Khanty-Mansi Autonomous Okrug vil. Lower Narynkar and vil. Peregrebnoe. He interviewed and, if possible, wrote down the glossary of the original vocabulary from 4 speakers of the Ob dialect of the Mansi language from the village Lower Narynkary: Butarina Galina Vasilievna, Kulikova Nina Iosifovna, Matveeva Zinaida Yakovlevna, Plekhanova Claudia Semyonovna, and from 2 native speakers from the village Peregrebnoye: Gyndysheva Taisiya Grigoryevna, Yarlina Evdokia Grigoryevna. Almost all speakers of the Ob dialect of the Mansi language are over sixty years old. As children, they spoke only the Mansi language, and did not know Russian before school. At present, at home with their children and grandchildren, they speak only Russian, communicate with each other in Mansi, but often switch to Russian.

The phonetic differences between the Ob dialect and the Mansi literary language are briefly described in [Rombandeeva 1973; Saynakhova 2012]. But they emphasised that, in general, the Ob dialect has not actually been studied. This is probably why researchers did not earlier note the presence of different-place accents in the Ob verb.

As a result of a complete analysis of the field material collected in 2013 by I. A. Stenin in the Praat program it was found that three verbal accent paradigms are distinguished in the Ob dialect of the Mansi language, see. [Normanskaya 2015b].

The specified material has already been processed and presented on the site <http://lingvodoc.ispras.ru/> in the form of dialect audio dictionaries interconnected by etymological links.

A comparative analysis of these dictionaries made it possible to significantly refine the Proto Mansian reconstruction. The first reconstruction of the Proto Mansian vocalism was proposed in [Kannisto 1919], who reconstructed 19 phonemes, for each of which another 5 to

about 10 variants of special reflections in one or several dialects were indicated without indications of systemic distributions. It is clear that if you simply take the series of correspondences, you will get not 19 phonemes, but more than 70. It was obvious to subsequent researchers that such a number of phonemes is unlikely, and, apparently, we are talking about phonetic variants. V. Steinitz, L. Khonti, M. A. Zhivlov, without referring to field audio material, proposed to reduce this system: V. Steinitz reconstructed 15 phonemes, L. Khonti and M. A. Zhivlov reconstructed 13 phonemes. But the correspondence between these systems is far from always unambiguous. The same phoneme in Kannisto, for example, \**a* can correspond to two different phonemes in later authors: \**ǎ*, \**ě*, and \**ę* even to three phonemes: \**ě*, \**ĩ*, \**ĩ̄*. And vice versa, \**ĩ̄* according to W. Steinitz corresponds to six phonemes according to A. Kannisto.

We suggested that in order to take research on the study of Mansi vocalism to a new level, it is necessary to verify the materials of A. Kannisto and B. Munkachi based on the material of modern dialects, and analyse them using the programs that experimental phonetics currently provides.

In the article [Normanskaya 2015c], a comparative analysis of field data on Mansi dialects was carried out and the Proto Mansi system of vocalism was reconstructed, consisting of 7 (? 8) phonemes: \**a*, \**ā*, \**u*, \**o*, \**i*, \**e*, \**ę*, ?, \**ĩ*.

In this article, we will conduct a comparative analysis of three existing sources of material on the Pelym dialect of the Mansi language: dictionaries [Slovtsov 1905], [MK 1986], [Kannisto 2013] and their comparison with the Proto Mansi reconstruction obtained in [Normanskaya 2015c], based on a comparison of field data on Mansi dialects.

Table 2  
Reflexes of Pro-Mansi vowels of the first syllable in the Pelym dialect

Proto Mansi	[Slovtsov 1905]	[Kannisto 2013]	[MK 1986]
* <i>a</i>	<i>a, я (C' _)</i>	<i>a</i>	<i>a/ā, ə</i>
* <i>ā</i>	<i>oa, #o, o#</i>	<i>ɔǎ</i>	<i>oǎ, ǎ (v_)</i>
* <i>i</i>	<i>u</i>	<i>i</i>	<i>i</i>
* <i>u</i>	<i>y, ю (C' _)</i>	<i>ulū</i>	<i>ulū</i>
* <i>o</i>	<i>o, ě (C' _)</i>	<i>ɔ, ǒ</i>	<i>ǎ, olō</i>
* <i>e</i>	<i>ə, e (C' _)</i>	<i>ě, elě</i>	<i>elě̄, elě̄</i>
* <i>ę</i>	<i>ə, e (C' _)</i>	<i>ě̄, ę</i>	<i>ā, əlā</i>

So, in conclusion of the comparison of data on Mansi dialects, we can conclude that the data of [Kannisto 2013] most closely correspond to the data of the Proto Mansi reconstruction made on the basis of a comparison of modern Mansi dialects. Each Proto Mansi phoneme has its own set of reflexes in the Pelym notes of A. Kannisto. It is noteworthy that not a single Pelym vowel grapheme by A. Kannisto is a reflex of two proto linguistic phonemes at once. But on the other hand, at present time, we have not been able to identify the positions of the appearances of several reflexes of one Proto Mansian phoneme in [Kannisto 2013], although PMans. \**o*, \**e*, \**ɛ* have two or more possible reflections: PMans. \**o* > *ɔ*, *ō*, \**e* > *ē*, *ē*, \**ɛ* > *ē*, *ɛ*.

In the dictionary of B. Munkacsy, the correspondences of the Proto-Mansi reconstruction are not so unambiguous. In particular, *ā* and *ə* in [MK 1986] may be Proto Mansi reflexes \**a*, \**ɛ*. Since they have special correspondences not only in the modern Mansi northern and eastern dialects, but also in the Pelym materials of A. Kannisto and Konstantin Slotvsov, then the distinction can be reliably postulated for Pelym, and the lack of distinction is a shortcoming of B. Munkacsy's recording of the Pelym materials. Presumably, this indistinguishability is due to the peculiarities of the native Hungarian dialect of B. Munkácsy, since, as the data collected on Hungarian dialects (see <http://lingvodoc.ispras.ru/>) show, for a native speaker of the Russian language *ɛ*, in many Hungarian dialects there is practically no difference from *a*. It can be noted that in the dictionary [MK 1986] almost all Pro-Mansch phonemes, with the exception of \**i*, have two or more reflexes (see Table 2). Their distribution positions could not be established either. It is interesting that A. Kannisto and B. Munkacsy's numerous reflexes of Proto Mansian phonemes in Pelym do not correlate with each other, that is, based on a comparison of the data from these dictionaries, no special series of correspondences are distinguished. It can be assumed that they reflect not phonological oppositions, but phonetic variants, which could vary according to individual dialects of each village or even for different speakers.

In the materials of Konstantin Slotvsov's those entries have a more phonological character. It is well known that the books prepared within the framework of the Translation Commission of St. Gury Kazansky, were aimed at native speakers and a simpler, more accessible record. If the Pro-Mansi phoneme has two entries in the Pelym dictionary [Slotvsov 1905], then the positions of their appearance are completely transparent: one of them is after palatalized consonants, on which the sign of palatalization is also in the materials of A. Kannisto and B. Munkacsy. In one case, as a comparison of the dictionary data [Slotvsov 1905] with the materials of A. Kannisto and B. Munkacsy and the data of modern dialects shows, one grapheme *e* (with the variant *e* after soft consonants) conveys two Proto Mansi at once phonemes \**e*, \**ɛ*. This is probably due to difficulties in conveying this difference with the symbols of the Cyrillic alphabet.

Thus, the comparison of the three dictionaries of the disappeared Pelym dialect shows that in the absence of external similarity in the rendering of individual words, all authors quite consistently and clearly reflected the Pelym vowel phonemes, but in the dictionary Konstantin Slotvsov presents an almost phonological record, and the dictionaries of A. Kannisto and B. Munkachi reflect pronunciation variants of vowel sounds. The Proto-Mansi reconstruction, made on the basis of a comparison of modern Mansi dialects, corresponds most exactly to the dictionary data [Kannisto 2013]. This corresponds to the traditional

opinion of specialists in the Ob-Ugric languages about the greater accuracy of the records of the Finnish linguist compared to B. Munkacsy.

But B. Munkacsy and Fr. Konstantin Slovtsov, there is an inaccuracy of transmission only for two PMans. phonemes: B. Munkacsy for PMans. \*a and \*ǣ, a y o. Konstantin Slovtsov for PMance. \*e and \*ǣ. It can be assumed that this is due to the peculiarities of the native languages of the compilers of dictionaries. It is noteworthy that the Russian priest Fr. Konstantin Slovtsov achieved an accuracy of rendering Pelym phonemes comparable to that of an outstanding Hungarian linguist, and it was the discovery of his dictionary and its introduction into scientific circulation that made it possible to verify the hypothesis of the highest accuracy of A. Kannisto's records.

### Abbreviations

K — Kondinsky dialects	T — Tavda dialects
KM — Middle Kondinsky dialect	VN — Northern Vagil dialect
LM — Middle Lozvinisky dialect	VNK — dialect of the village Kama
LO — Upper Nelozvinisky dialect	
LU — Lower Nelozvinisky dialect	
N — North Mansi dialect	
P — Pelym dialect	
So — Sosva dialect	
Sy — Sygvin dialect	

### Literature

Zhivlov 2006 — Zhivlov M. A. Reconstruction of the Proto-Slavic vocalism.

Kuzakova 1963 — Kuzakova E. A. South Mansi (Kondinsky) dialect (in comparison with the North Mansi dialect). L., 1963.

Normanskaya 2015a — Normanskaya Yu. V. The system of multi-place stress in the Mansi verb and its external parallels. Part I. The Ob dialect of the Mansi language // Ural-Altai studies. 2015, 2 (17). pp. 51-66.

Normanskaya 2015b — Normanskaya Yu. V. The system of multi-place stress in the Mansi verb and its external parallels. Part II. Yukondinsky dialect of the Mansi language // Ural-Altai studies. 2015, 3 (18). pp. 88-103.

Normanskaya 2015b — Normanskaya Yu.V. New field and archival data on Mansi dialects and their significance for the Pramanian reconstruction of the system of vocalism of the first syllable // *Ural-Altai studies*. 2015, 4 (19). pp. 40-59.

Rombandeeva 1973 — Rombandeeva E. I. Mansi (Vogul) language. Phonetics. Morphology. Word formation. M., 1973.

Saynakhova 2012 — Saynakhova A. I. Dialectology of the Mansi language

Slovtsov 1905 — Experience of the Russian-Vogul dictionary and translations into the Vogul language (compiled by the priest of the Verkhne-Pelymsky (Spassky) parish, Fr. Constant. Slovtsov). Tobolsk, 1905.

Honti 1982 — *Honti L.* Geschichte des obugrischen Vokalismus der ersten Silbe. Budapest, 1982.

Kannisto 2013 — *Kannisto A., Eiras V., Moisio A.* Wogulisches Wörterbuch. Helsinki: Societé Finno-Ougrienne: Kotimaisten Kielten Keskus, 2013.

MK 1986 — *Munkácsi B., Kálmán B.* Wogulisches Wörterbuch / Gesammelt von *Munkácsi B.* Geordnet, bearb. und hrsg. von *Kálmán B.* Budapest, 1986.

Steinitz 1955 — *Steinitz W.* Geschichte des wogulischen Vokalismus. Berlin, 1955.

---

<sup>1</sup> The work was done in assistance of grants **PHΦ** № № 20-18-00403.