Uralic studies, languages, and researchers

Edited by Sándor Szeverényi

Studia uralo-altaica 54

Uralic studies, languages, and researchers

Proceedings of the 5th Mikola Conference
19–20, September 2019

Edited by Sándor Szeverényi

Szeged, 2021

Table of contents

Undiscovered treasures: From the field research archive to the digital database

Beáta Wagner-Nagy, Chris Lasse Däbritz, and Timm Lehmberg

Universität Hamburg

1. Introduction[1]

Finno-Ugric Studies, or to be more precise, Uralic Studies traditionally belongs to the field of empirical linguistics, even if —especially in the past— the collected data were often processed in the context of historical linguistics. It must be noted that apart from some languages, such as Hungarian, Komi, or Finnish, which were written relatively early, most Uralic languages have no —or only very few— historical written sources. Of course, this does not mean that we do not have any archival data from these languages that are older than the written form in these languages. Even if the data are recent in comparison to e.g. Hungarian, Greek, or Latin, we still have data. Nevertheless, these data were not created naturalistically, that is, they are not written thoughts or literary pieces, and they are not even translations from other languages, etc. but were created as researchers' field records. For example, the older records by Gerhard Friedrich Müller, or, somewhat later, by Matthias Castrén or Kai Donner, are records that the researchers wrote down, and not records of the respective speakers studied by them. In the records made later, such as the manuscript materials of Angeline Kuzmina in the 1960s, one can also find recordings that were not made by the researcher but by a speaker themselves. All these materials were created during or after field research and, usually, only a small portion of the collected materials have ever been published. On the one hand, researchers usually collect much more data than they then use for their publications, such as the grammars and dictionaries of the Samoyedic languages in the case of Castrén (1854, 1855). On the other hand, in some cases, as in the case of Kuzmina's notes, very few records were published (see Tuchkova and Helimski 2010), or even as in the case of Müller's notes, only small parts of the work were used by other researchers. However, access to primary data

may also be important for future generations of researchers. The processing of unpublished materials can bring new knowledge about languages or peoples, but as Bowern (2008: 185) has pointed out, studying the archive materials can also bring knowledge about the way researchers work.

This study aims to throw some light on the special nature and potential of these resources and to open the discussion for further issues of, beyond other things, information modeling, and visualization. Section 2 focuses on what archives are and how to find archive materials, such as field notes of researchers. The first thing to think about in this topic is, however, what field notes are. This topic is discussed in section 3. In section 4, two examples are provided, specifically about how we can open field notes in a modern way.

2. Archives and archive materials

It is common in science and scholarship, especially in the early times when the main part of the work was analog rather than digital, that after the death of a researcher their (scientific and/or scholarly) material was preserved by their successors or family. But one question arises: where? The simplest answer is: in an archive. But if this is the answer to this question, we must first clarify what an archive is. The easiest answer is that an archive is a physical place where letters, reports, notes, memos, photographs, drafts, and final manuscripts as well as other primary sources are kept. Assuming that we accept this simple definition, we can name many places archives. And so we do! For example, an archive is often referred to as an individual cupboard or drawer in which, for example, field note manuscripts are kept, without this place ever being an official archive. This can be called a private archive. Some archives are at least in an institution, others are actually in private households. Some of them may be cataloged and ordered, others are not. Only a few field notes are preserved in official archives. However, archives of private companies or laboratory archives as well as those of associations, unions, and political parties are also private archives. In contrast, public archives are usually maintained by public services, such as administrative institutions, governments, etc. But the main task is the same: "to systematically take over, record, order, permanently store and index written, image and sound carriers and electronic storage media from public services, other institutions (associations, companies) or individuals" (Reimann 2004: 20). Examples of storing field research materials in a public archive can also be found, e.g. the field notes by Gerhard Friedrich Müller (e.g. section 4.1 below).

Everything that is kept in an archive is called archive material (German *Archivgut*). However, typical archive materials are documents of administrative activities, letters, etc. These materials are usually unique, their copies or reproductions can never replace the original. This also applies to the field notes, even if despite their variation and

uniqueness, they are extremely rarely recognized as research objects but as a helpful tool to gain relevant data. Hence, they are not curated and preserved as other empirical material is, not even if they can be considered a manuscript.

3. What are field notes and where do field notes come from?

Every researcher who does fieldwork, be it linguistic or ethnographic-anthropological, takes memos, notes, etc. We can call these notes *field notes*. Field notes do not represent the object of research itself, they represent the researcher's reception of the object of research. This holds also for drawings of artifacts in archeology and descriptions of something observed, whereby they usually contain manifold and multi-layered information of interdisciplinary relevance. Due to the individual character of field notes, there is no standard format and content for these records. Depending on the objective or the goal of the target work, these data may be primary language materials, but they may be notes about the people or their culture. If we look at it from the perspective of a linguist, we can and must conclude that beyond the primary data, the data contain secondary linguistic materials as well. Such information may include, for example, information about the target language, translations, or analyses of the language data and geographical information as well as metadata of individuals and personal notes. Thus, field notes always contain multi-layered information of interdisciplinary relevance, and often in multilingual form.

The primacy of the language data can already be debated in so far as it always is a written representation of *eo ipso* oral material. This has two important consequences: first, field notes always and necessarily contain an interpretation of the researcher who wrote them down, and, second, field notes contain more information than just the object of research itself. To illustrate this, the image below (cf. Figure 1) shows an instance of linguistic field notes, coming from the Kuzmina Archive, which is stored at the Institute of Finno-Ugric/Uralic Studies at the University of Hamburg. It can be seen at first glance that the researcher uses at least two languages (Selkup and Russian). The word-by-word glosses are given in Russian. In addition to the Selkup words, one can find corrections written in a different pen. It is not indicated who made these corrections and when, but from the handwriting, we can deduce that it is the same person who made the field notes.
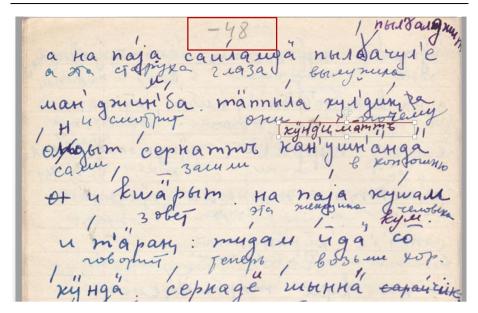
Figure 1. Manuscript field note page from Kuzmina Archive
(volume 26, book 2, page 48)

As mentioned above, many of these records are never researched. There are several reasons for this. One of them is the fact that field notes are hardly ever considered as a research topic; however, they are unique written artifacts. Another reason is that researchers always work on only a part of the collected material. A linguist publishes the linguistic data, such as texts or elicited sentences, lexical materials, etc., and probably the associated metadata as well as describes the language based on the data collected, therefore interpreting the linguistic material (cf. Himmelmann 2006, Woodbury 2011, among others). An anthropologist looks at the data from his or her point of view. A further aspect is that many field notes are either partially or not at all accessible to other researchers or the broader audience.

From the history of Uralic studies, we know very well that the collected materials did not always find their way to the scientific community. There are, again, several reasons for this. Often the researchers were not able to publish their materials, for instance, because of their untimely death (e.g. Castrén or Donner, among others) or for other, sometimes political reasons (e.g. Gerhard Friedrich Müller). Other researchers then edited the materials from their estates as in the case of Donner, whose Kamas materials were compiled and published by Aulis Joki, or in the case of Castrén, whose fieldwork materials were compiled, edited, and published by his friend Anton Schiefner. Nevertheless, the question of what happens to the original recordings keeps arising. This issue is considered in section 4 below. And as noted above, it would

often be useful to be able to look at the original materials as well, and not just the final product such as a published text collection or grammar.

4. Publishing old fieldwork materials

As already mentioned, the goal of the practice of working on manuscript field notes seems to be to extract the data of the given researcher's interest and publish it. This is what happens with many field note manuscripts. Either the linguistic data is extracted and, for example, a grammar is compiled or only the collected texts are published. However, ignoring manuscript field notes as research objects can lead to misunderstandings and even a loss of quality in any further research based on them. An example may illustrate this: the Finnish scholar Matthias Alexander Castrén conducted several field trips to numerous places in Northern Eurasia in the first half of the 19th century. Passing away at the age of only 38, he left a lot of materials and data from his trips, e.g. detailed notes on the grammar of all Samoyedic languages. In 1854, Anton Schiefner compiled a comparative grammar of the Samoyedic languages on the basis of Castrén's field notes (Castrén 1854). Without denying the merits of Schiefner's work, it must nevertheless be noted that Castrén's field notes are much more detailed and provide much more information (cf. Däbritz unpublished). However, Schiefner's work never aimed at researching the field notes as manuscripts or as written artifacts, instead, he wanted to publish the first results on these languages. With today's technical capabilities, however, it is possible to make the manuscripts accessible and available for the research community and the general audience so they can be explored in themselves as manuscripts.

However, this raises the question, how? It must be borne in mind that the general audience and the specialists sometimes have completely different needs. The options will be explained via two examples. The first is Gerhard Friedrich Müller's collection, the second is Angelina Kuzmina's field materials.

4.1. Müller's materials

Gerhard Friedrich Müller (1705–1783) was a historian and ethnographer who was invited to Sankt Petersburg to co-found the Russian Imperial Academy of Sciences. He not only participated in the Second Kamchatka (also known as Great Siberian) Expedition (1733–1743), but he was the leader of the "scientific group". During the field trip, he described the people he met and collected not only geographical and historical but extensive linguistic materials as well. He has collected words of semantic fields from various languages. These included numerals, religious vocabularies, such as the names of god and devil. Müller used a word list with approximately 240 lexemes. The source language was mostly Latin, sometimes

Russian. The notation of what he used can be considered phonetic. After returning from the expedition, Müller published only some ethnographic materials, among other things the book titled *Report on three pagan peoples, the Cheremis, Chuvash, and Votiak, living in the vicinity of the city of Kazan* (1759). He devoted his time to the study of the history of Siberia. Thus, not only a good portion of his ethnographic descriptions but also his linguistic materials remained unpublished for centuries. From the 1980s onwards his ethnographic records, letters, and other materials were gradually published (cf. Bucher 2002, Hoffmann 1995, Müller 2010, 2018 among other publications; for the Kamchatka Expedition, see Dahlmann 2009). But this cannot be said about the linguistic materials: they remain unpublished still. However, this does not mean that these materials were never seen by anybody. Helimski has dealt with Müller's estate several times. He not only described Müller's manuscripts preserved in the Russian State Archives of Ancient Acts (RGADA)[2] in Moscow and other archive materials from the 18th century (Helimski 1993, 1987) but also used Müller's materials in his several publications. However, Helimski never published the whole material.

Before we consider how these legacy materials could be published or presented, let us take a look at Müller's manuscripts. What materials do we have? How is the data organized, how is it written? Above all, where exactly are they located? This latter question can be answered quite simply. All those who are familiar with the history of research or who know the works on Müller know where these materials are preserved or know in which article or book the corresponding information can be found. As mentioned above, these materials are in the Russian State Archives of Ancient Acts. And here they are located in archival collection 199, archival inventory 2, portfolio 513, files 2 and 7.[3] Finding this collection is not difficult. The archive has an electronic catalog where one can easily find the collection. Figure 2 shows the search results.

---

[2] http://www.rgada.info/

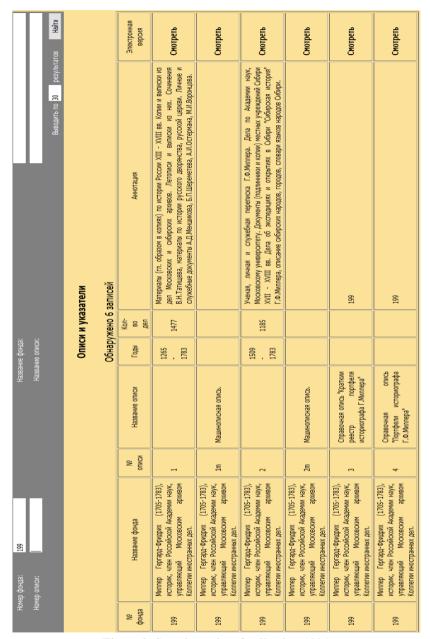[3] фонд архивного собрания 199, опись архивного фонда 2, портфель 513, дело 2, 7.

Номер фонда: 199
Номер описи:
Название фонда:
Название описи:

Выводить по 30 результатов   Найти

Описи и указатели

Обнаружено 6 записей

№ фонда	Название фонда	№ описи	Название описи	Годы	Кол-во дел	Аннотация	Электронная версия
199	Миллер Герард-Фридрих (1705-1783), историк, член Российской Академии наук, управляющий Московским архивом Коллегии иностранных дел.	1		1265 - 1783	1477	Материалы (гл. образом в копиях) по истории России XIII - XVIII вв. Копии и выписки из дел Московских и сибирских архивов. Летописи и выписки из них. Сочинения В.Н.Татищева, материалы по истории русского дворянства, русской церкви. Личные и служебные документы А.Д.Меншикова, Б.П.Шереметева, А.И.Остермана, М.И.Воронцова.	Смотреть
199	Миллер Герард-Фридрих (1705-1783), историк, член Российской Академии наук, управляющий Московским архивом Коллегии иностранных дел.	1m	Машинописная опись.				Смотреть
199	Миллер Герард-Фридрих (1705-1783), историк, член Российской Академии наук, управляющий Московским архивом Коллегии иностранных дел.	2		1509 - 1783	1185	Ученая, личная и служебная переписка Г.Ф.Миллера. Дела по Академии наук, Московскому университету. Документы (подлинники и копии) местных учреждений Сибири XVII - XVIII вв. Дела об экспедициях и открытиях в Сибири "Сибирская история" Г.Ф.Миллера, описание сибирских народов, городов, словари языков народов Сибири.	Смотреть
199	Миллер Герард-Фридрих (1705-1783), историк, член Российской Академии наук, управляющий Московским архивом Коллегии иностранных дел.	2m	Машинописная опись.				Смотреть
199	Миллер Герард-Фридрих (1705-1783), историк, член Российской Академии наук, управляющий Московским архивом Коллегии иностранных дел.	3	Справочная опись "Краткий реестр портфеля историографа Г.Миллера"			199	Смотреть
199	Миллер Герард-Фридрих (1705-1783), историк, член Российской Академии наук, управляющий Московским архивом Коллегии иностранных дел.	4	Справочная опись "Портфели историографа Г.Ф.Миллера"			199	Смотреть

Figure 2. Search results of collection 1994

If we look at these results, we notice that we have not received any results for file 7. This means that the cataloging has not yet been completed. Some manuscripts, however, are already available on the site. If we click on inventory 2, we get to the digitized manuscripts. Unfortunately, the digitized version does not contain portfolio 513. However, for linguistic research, this portfolio would be interesting. The researcher then has to visit the archive.

Let's look at the materials anyway. What can be said about the physical material? We know that at Müller's time, the paper was naturally hand-made, from textile waste (rags). This paper is much thicker than the conventional paper used today and has various sizes. We can assume that Müller used loose pieces of paper, which were bound together with adhesive tape at a later point in time. When exactly this happened is not known, but from its appearance, we can assume that it did not happen in Müller's lifetime but much later. Figure 3 shows a piece of the manuscript. In the middle of the picture, the tape is visible.
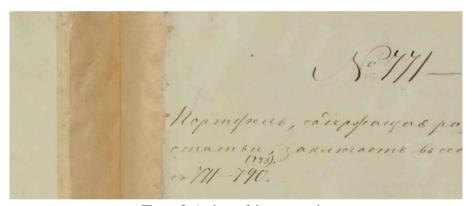

Figure 3. A piece of the manuscript
(Source: http://rgada.info/opisi/199-opis_2/0260.jpg; Accessed on 18 September 2019)

As Helimski already described (1987: 2), specialists know significant parts of these materials. In various works various parts were published, e.g. Helimski published the Mator and Kamas wordlists in the work mentioned above. He not only published the list but also commented on the materials and attached a photocopy of the manuscript. Would it be worth working with these manuscripts again? Most likely, yes. When we look at these manuscripts, we immediately see that some parts were already prepared for publication. The words are already organized by language and semantic fields. If we decide to publish the whole of the manuscript, we would have to clarify in what form it should be published. Should it be in paper form, as in Helimski's work, or in digital form? One graph-based approach to be mentioned in this context that takes into

account both the multi-layered nature and the individual character of field note collections is described by Jettka and Lehmberg (2020). In this case, connections between the semantic categories could be established relatively quickly. To do this, a set of prerequisite sub-objectives is to be accomplished, such as (a) the digitization of the wealth of analog archived materials, (b) the conversion of the digitized material into a searchable electronic database that would allow multiple modes of analysis and information extraction, and (c) the development of the capacity for dynamic visualizations of analytical results. A possible result would be, for example, as illustrated in figure 4. Here the search query was Kamas, and only Kamas items were searched for.[5]
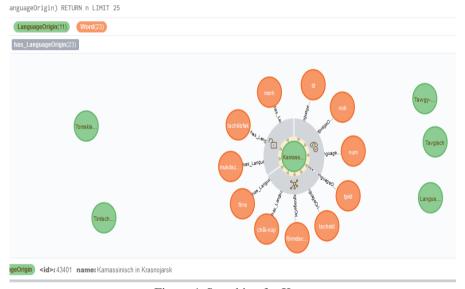

Figure 4. Searching for Kamas

A somewhat more complex search is shown in Figure 5. The search term 'windstorm' (blue ball) was used. The green units (balls) show the languages in the Müller Archive (red ball), the brown ones are the common names. The results show that the searched word is *tschóbfak* in Kamas. The graph also shows the relations. The word comes from Müller's archive, in the same archive, other units can be found, which are also presented and connected to another language. One can also see that there are more Kamas words to be found in Müller's collection.

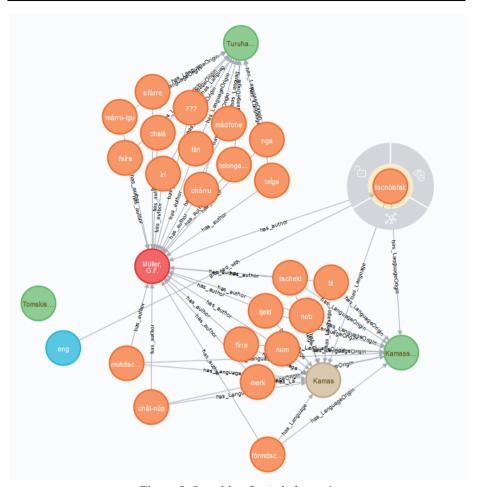[5] The search results do not represent the entire material, but only show a working version.

Figure 5. Searching for 'windstorm'

Of course, beyond simply searching for lexemes and the links between them, one can also create statistics. If one sorts Müller's language data according to today's names or puts the names together, one can visualize a relationship. This is shown in Figure 6. Note that the figure does not show the actual relation in Müller's manuscript, but the current processing status in the database.[6]

---

[6] The data for graphic search and for the statistics were prepared by Daniel Jettka, for which the authors are very grateful.
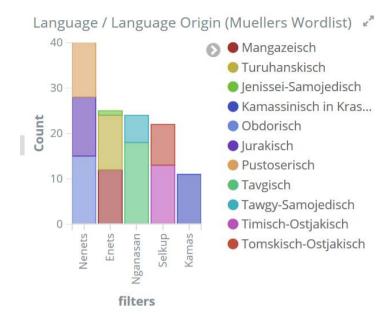
Figure 6. Visualization of the statistics of Müller's wordlist

Such a publication of the data shows well the relations between the collected lexemes and allows for a quick generating of statistical analyses. In this format, it is difficult to write, for instance, comments on the individual lemmas. For this reason, it cannot necessarily replace the traditional paper-based publication, but it can complement it. However, the original manuscript is still not accessible.

4.2. Kuzmina's materials

Angelina Kuzmina's fieldwork manuscripts are preserved at the Institute of Finno-Ugric/Uralic Studies[7] in Hamburg. Kuzmina was a student of Dulson, who worked in Tomsk on Siberian languages. She researched the German dialects of Siberia before turning to the study of the Selkup language. Between 1962 and 1969 and later in 1976 and 1977 she visited all regions where the Selkups lived and collected valuable materials. Her estate consists of 35 handwritten volumes, 30 of which are preserved in Hamburg. The volumes contain 357 books, with a total of 8,554 pages. When working with these materials, several details must be taken into account. Although an initial review of the materials took place and a temporary catalog was published (Tuchkova and Helimski 2010), the materials were not fully cataloged. Consequently,

---

[7] https://inel.corpora.uni-hamburg.de/portal/kuzmina (Accessed on 26 November 2020).

cataloging is necessary so that the wider audience and the research community can find out what is in the field notes. A very simple solution for this is a tabular view, but this offers few options. For the first search, however, it is ideal.[8] The second issue, however, is the original question of how to make the manuscripts and content available as research objects. Figure 7 shows the cover of one field notebook.
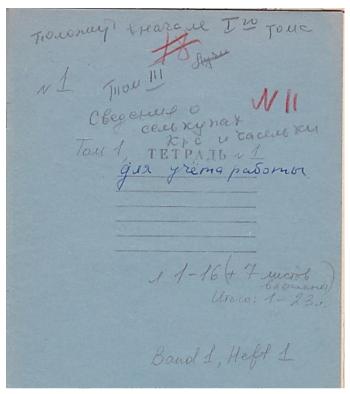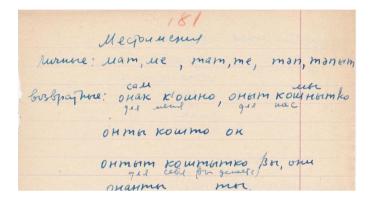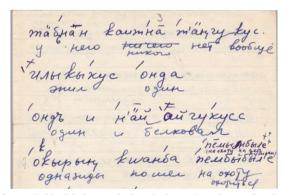

Figure 7. Cover of the first notebook in volume 1

At first glance, one can see that there is a lot of information available here. It is indicated which notebook of which volume this is (notebook 1 in volume 3). There is additional information, which was most likely created later and contains instructions and remarks. It is also visible that a later numbering with a red pencil was also done (Nr 11). And a third numbering occurred even later. This is at the bottom of the cover and was made by Eugen Helimski. It is not only his handwriting that indicates this,

---

[8] See Lehmberg 2020 for the potential that arises from the application of data analysis and visualization routines to a linear structured information resource derived from the Kuzmina Archive.

but that this numbering is in German. Nevertheless, the question arises as to who the red numbering comes from. It would be easy to say that they come from Kuzmina. This red numbering is not only on the cover, but the pages are numbered in some booklets also with a red pencil. However, this holds true only for the booklets that originate from the earlier research trips, but not for those that were written in the late 60s or the 70s: there the numbering is written in grey pencil. Compare the two pages below in Figures 8a and 8b.

Figures 8a and 8b. Volume 2, book 9, and volume 26, book 1.

We know that Kuzmina left Tomsk in 1964 and moved to Novosibirsk, where she worked until the end of her life. If we open the field research manuscripts which are preserved in Tomsk and were written in Dulson's lifetime, we find exactly such red numbers as in Kuzmina's manuscript; in the notebooks written by Dulson as well as in the notebooks prepared by his students. All these facts indicate that these red numbers were made by Dulson.

As mentioned earlier, the field notes contain not only linguistic data but also other materials, such as metadata, sketches, and maps. This is true of Kuzmina's materials as well, which contain maps, metadata about the speakers, elicited sentences, conjugated verb forms, etc. During the processing of the fieldwork materials, the collected texts are usually published. It can be done as an electronic resource, such as a transcript in a corpus. But it can also simply be published in paper form. During Kuzmina's lifetime, only the latter option existed, so she published some texts in text collections. Regardless of which publication method we choose, some of the information contained in the field notes or which derivable from them remains hidden. The *INEL Selkup Corpus* (Brykina et al. 2020)[9] contains a text (AGS_1964_HeroAndDragon1_flk) which was published by Kuzmina herself. This information is given in the corpus, in the metadata of the text. The information in which volume this text can be found is also given. While editing the text, a transcript is generated and added to the corpus. Even if the INEL Selkup Corpus contains a relatively large amount of information, such as the original Cyrillic spelling, some corrections, etc., some properties remain hidden. The reason is that a corpus or a transcript in a corpus can never reproduce everything from a manuscript. Figure 1 above shows a part of the manuscript. Figure 9 shows the beginning of the text.
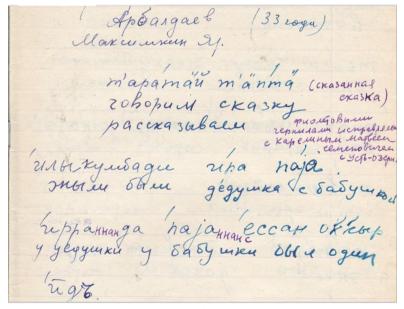

Figure 9. First sentences of the text in the manuscript

Here we get the information about not only who was recorded providing the text, but we will also explain who made the purple corrections. Kuzmina checked the text with another Selkup speaker. Looking at the metadata, it can be seen that the speakers come from two different villages. The storyteller (Arbaldaev) lived in Maksimin Yar, the other speaker, Karelin Matveev Semenovich, lived in Ust-Ozyornoe. Even if they speak the same dialect (Middle Ket), they have pronunciation variants and possible grammatical differences as well. Right at the beginning of the text, we see an important correction by Karelin. The second sentence is a possessive construction, in which normally the possessor is marked with the adessive suffix, but in the field notes, the possessor is unmarked. Karelin corrected it, of course. Well, two questions must be asked: is the possessor unmarked in the variety spoken by Arbaldaev? In this case, there is a dialect difference, or at least an idiolectal difference. This question can be answered very quickly by examining the other texts supplied by Arbaldaev. In all other texts the possessor was marked by him also with a postposition, so, obviously, here there is a mistake. It is not possible to find out who the error came from. It can be a recording error by Kuzmina, but it can also be an error by Arbaldaev. But the second question is, how should the text be published: with or without this grammatical mistake. From the metadata, one can know that this text was already published by Kuzmina (1967: 294–296, 317–318). The published version is shown in Figure 10.
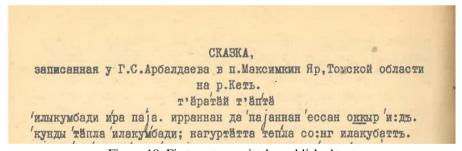

Figure 10. First sentences in the published text

Here we can see that Kuzmina has incorporated the corrections and corrected the ending. Concerning this text, there are two more records among the manuscripts. The second version of the text is in book 4 (88–96) in the same volume. This version was written by Arbaldaev himself, not by Kuzmina. The text is not identical to the previous one (book 2, page 25), but the content is the same. As figure 11 shows, the correction is still missing. Accordingly, another variant must be considered.
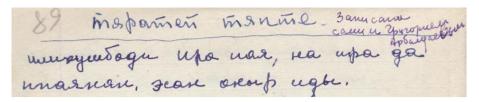
Figure 11. The first sentences of the text written by Arbaldaev

The third variant is an envelope in volume 26, which contains a typed copy of the text from 1964. As figure 12 shows, the correction is already incorporated. So it must have been made between 1963 and 1964. It is not known when exactly, but with a closer examination of the data in the archive, it could be found out. For this, it would be necessary to evaluate when Kuzmina worked with Kalinin. Most of his material is from 1963, but obviously, there was a later point in time when Kuzmina consulted with him. This still requires further research and evaluation of the materials.
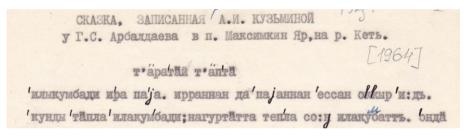

Figure 12. The first sentences of the text in the emvelop

5. Summary

From the examples discussed above, it can be seen that a closer look at the field research materials can provide a lot of information that can explain not only the origin of the texts but also the working methods of the researchers or the speakers they worked with. All this information is part of the text that is likely to be published. For this reason, when working with such archives as the Kuzmina archival materials, one has to think about how to make what is derived from the manuscripts accessible. A good approach is what the project chose: the manuscripts and the published versions of all texts, if they are available, are published. But one can go one step further and build the whole archive catalog in such a way that one links the geodata, the catalog units, and the transcripts that are in the corpus. This would make it easier to find the information. Future approaches that aim at the sustainable (digital) availability of all information contained in these resources will have to find both resource-specific and generic way of information modeling and visualization.

References

Bowern, Claire 2008. *Linguistic fieldwork: A practical guide*. London: Palgrave Macmillan Limited.

Brykina, Maria, Svetlana Orlova, and Beáta Wagner-Nagy. 2020. INEL Selkup Corpus, Version 1.0. Publication date 2020-06-30. Archived in Hamburger Zentrum für Sprachkorpora. http://hdl.handle.net/11022/0000-0007-E1D5-A. In: Wagner-Nagy, Beáta, Alexandre Arkhipov, Anne Ferger, Daniel Jettka, and Timm Lehmberg (eds.). *The INEL corpora of indigenous Northern Eurasian languages.*

Bucher, Gudrun 2002. Von Beschreibung der Sitten und Gebräuche der Völcker: die Instruktionen Gerhard Friedrich Müllers und ihre Bedeutung für die Geschichte der Ethnologie und der Geschichtswissenschaft. Stuttgart: Steiner.

Castrén, M. A. 1854. *Grammatik der samojedischen Sprachen*. Herausgegeben von Anton Schiefner. St. Petersburg: Kaiserliche Akademie der Wissenschaften.

Castrén, M. A. 1855. *Wörterverzeichnisse aus den samojedischen Sprachen*. Bearbeitet von Anton Schiefner. St. Petersburg: Kaiserliche Akademie der Wissenschaften.

Dahlmann, Dittmar 2009. *Sibirien vom 16. Jahrhundert bis zur Gegenwart*. Paderborn – München – Wien – Zürich: Ferdinand Schöningh.

Däbritz, Chris Lasse. *Manuscript fieldnotes as multilayered written artefacts: An interdisciplinary approach*. Unpublished manuscript.

Helimski, Eugen 1987. Two Mator-Taigi-Karagas vocabularies from the 18th century. *Journal de la Société Finno-Ougrienne,* 81: 49–132.

Helimski, Eugen 1993. Samoyedic vocabularies from the 18. century: A list of archive manuscripts. *Ural-Altaische Jahrbücher: NeueFolge,* 12: 249–265.

Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In: Gippert, Jost, Nikolaus P. Himmelmann, and Ulrike Mosel (eds.): *Essentials of language documentation*. Berlin: de Gruyter, 1–31.

Hoffmann, Peter (ed.) 1995. *Geographie, Geschichte und Bildungswesen in Rußland und Deutschland im 18. Jahrhundert: Briefwechsel Anton Friedrich Büsching – Gerhard Friedrich Müller 1751 bis 1783.* Berlin: Akademie-Verlag

Jettka, Daniel and Timm Lehmberg 2020. Towards flexible cross-resource exploitation of heterogenous language documentation data. In: *Proceedings of the 12th Conference on Language Resources and Evaluation* (LREC 2020), 2901–2905. available at https://www.aclweb.org/anthology/2020.lrec-1.354.pdf (Accessed on 17 July, 2020)

Kuzmina, Angelina [Кузьмина, А.И.] 1967. Диалектологические материалы по селькупскому языку. In *Исследования по языку и фольклору*. Новосибирск, 267–329.

Lehmberg, Timm 2020. Digitale Edition des Kuzmina Archivs. *Finnisch-Ugrische Mitteilungen* 44: 121–131.

Müller, Gerhard Fridrich 2010. *Ethnographische Schriften* 1. Edited by Wieland Hitzsche. Halle: Verlag der Frankeschen Stiftungen.

Müller, Gerhard Fridrich 2018. *Ethnographische Schriften* 2. Edited by Wieland Hitzsche and Aleksandr Elart. Halle: Verlag der Frankeschen Stiftungen.

Reinmann, Norbert 2004. Grundfragen und Organisation des Archivwesens. In: Reinmann, Norbert (ed). *Praktische Archivkunde*. Münster: Andrey Verlag.

Tuchkova, Natalia and Eugen Helimski [Тучкова, Н.А. and Хелимский, Е.А.] 2010. *О материалах А.И. Кузьминой по селькупскому языку = Über die selkupischen Sprachmaterialien von Angelina I. Kuz'mina*. Institut für Finnougristik/Uralistik der Universität Hamburg.

Woodbury, Anthony C. 2011. Language documentation. In: Austin, Peter K. and Julia Sallabank (eds.): *The Cambridge handbook of endangered languages*. Cambridge: Cambridge University Press. 159–186.