

The Use of Archives and Fieldwork for the Study of the Endangered Languages of Russia

Tjeerd de Graaf

Department of Linguistics
Groningen University
degraaf@let.rug.nl

Abstract

This contribution considers some collaborative projects of Groningen University with Russian institutions. A survey is given of the use of sound archives in St. Petersburg with large collections on the languages of Russia. In the framework of our projects part of this material has been recovered, such as recordings made of colonial German dialects, which are now available in an acoustic database. This is also the case for local Russian dialects, the Yiddish language and other language groups. The material is used for a further study of language contact, folklore and other aspects of Russia's cultural heritage. This historical material is complemented by the results of present-day fieldwork in various regions of Russia. Specific results on the language of the Siberian Mennonites and Yiddish are mentioned, whereas the study of language contact in the Russian North and phonological properties of Siberian languages are also current research projects. The results of these projects make it possible to learn about the cultures and languages of Russia and they provide new methods of teaching these topics, not only in universities, but also in special locations for the minority peoples themselves.

1. Introduction

Traditionally research institutes in the Russian Federation have paid much attention to the multicultural aspects of the Russian society, and in the last ten years interest in these fields is again growing. This holds in particular for the University of St. Petersburg, where in the past many scholars have been active in the fields of language description, ethnolinguistics and related disciplines involving the languages of Russia. Large collections of data on these languages and cultures can be found in the archives and museums of St. Petersburg. After a description of some early Dutch contributions in the initial development of Russian science and education, we shall mention some recent projects, which we have undertaken in collaboration with the University and the Russian Academy of Sciences in Saint-Petersburg. This collaboration has been financially supported by the EU organisation INTAS and the Netherlands Organisation for Scientific Research. In our approach we reconstruct sound recordings of endangered languages and cultures and create databases of these materials. The historical data from these archives are supplemented by the results of fieldwork expeditions. In 1990, 1991 and 2000 we took part in ethnolinguistic fieldwork on the island of Sakhalin (De Graaf, 1992), and in 1994 we joined expeditions to the Arctic Peoples near the Kolyma River in Northern Yakutia (De Graaf, 1996). In the Altai Region of Southern Siberia we visited ethnic minorities, such as the Siberian Mennonites (De Graaf and Nieuweboer, 1994). In this way we investigated processes of language contact, language shift and language death for the different ethnic groups of Russia.

2. Early scientific links: Nicolaas Witsen.

During a stay in the Republic of Sakha (Yakutia) in 1994, local linguists in Yakutsk told us about the history of the Yakut language. They mentioned the fact that the first written information on this language could be found in a

book by the Dutch author Nicolaas Witsen, but that they were not yet able to read this text. Witsen was one of the Dutch people who had personal contacts with Peter the Great during his visit to Amsterdam in 1697 and who provided information about Western Europe to the Russians. At the same time the Western world learned much about the Russian Empire and its inhabitants from Witsen's publications. One of his books, *De Moscovische Reyse*, about a trip he made to Moscow (Witsen, 1666) and about his stay there in 1664-65, was translated into Russian and published in 1996 (Witsen, 1966).

A more important work is contained in two large volumes of the book *Noord en Oost Tartarye* which first appeared in 1692 with a description of the Eastern Parts of the Russian Empire. In this book Witsen gives many details on the peoples of Russia, their languages and cultures, and he provides the first maps of this part of the world. For many of the Siberian languages, for example for Yakut, word lists are provided. The books are written in seventeenth century Dutch and it is very difficult for colleagues in Russia to get access to the interesting material it contains, such as the linguistic data on the Yakut language.

In the framework of the Nicolaas Witsen Project, an interdisciplinary team of specialists has been created for the investigation of the minor peoples of Northeast Asia and their cultures which were described by Nicolaas Witsen. With this aim we established contacts with foreign institutions involved in the study of Siberia and the Russian Far East, such as those in Russia (Saint-Petersburg and Moscow) and in Japan. They represent the various aspects of this project: history, ethnology, linguistics, geography, etc. Research has been initiated in (mainly Russian) archives in order to find out where Witsen got his information. Much of the linguistic material has been investigated and compared to the present day language situation in Siberia. The minorities in the Asian part of Russia are further studied by organising fieldwork expeditions and the results of these investigations will be published in separate articles and

books. It is important to pay attention to the survival of the language, culture and oral traditions of these minorities and to preserve as much material as possible by creating a database of this cultural heritage. For this purpose modern techniques of information technology are applied. The results of this international collaboration will be interesting for specialists in various fields, such as history, ethnology and linguistics, and it will provide many important data from Russian archives, which are relevant for further international research projects.

3. Sound Archives; the INTAS projects

Archives contain not only written material, but also other data such as sound recordings. These recordings are stored in sound archives all over the world. Many of them can be found in the archives of the Russian Academy of Sciences, which is part of the Pushkinsky Dom (Pushkin House), the Russian Museum of Literature in Saint-Petersburg. These recordings contain valuable information on the cultural developments in the Russian Federation, in particular spoken language and songs recorded at the beginning of the 20th century. Together with the Department of Phonetics of Saint-Petersburg University we have restored these recordings and put them into a database, which is now available for all colleagues who can use the material for various purposes. This project on *The Use of Acoustic Data Bases and the Study of Language Change* was financially supported by the organization INTAS in Brussels (De Graaf, 1998a).

INTAS stands for INternational ASsociation for the Promotion of Co-operation with Scientists from the Independent States of the former Soviet Union. This organization is part of the European Union in Brussels and it plays a very important role in the exchange of scientists between the Russian Federation and Western Europe. Background information on our INTAS projects is given here in order to illustrate some useful results which have been obtained in this exchange program. Our aim is to reconstruct more old recordings of the Russian Academy of Sciences in the Pushkinsky Dom and to use these data for the study of language, music and folklore of the peoples of Russia (De Graaf, 1997).

Prior to 1890, linguistic and ethnological fieldwork was based on direct contacts with representatives of various cultures, in which the investigator took notes by hand after many repetitions of tales and songs during recording sessions. This was a laborious process for both the investigator and the performer or informant. At the end of the nineteenth century, the great invention of the phonograph by Thomas Edison changed all this. For the first time in human history, people were able to store and rehear acoustic data, in particular speech, songs and music. As recordings were made, it became obvious that a central facility was needed for the preservation of the valuable material, which had been collected. At the beginning of the twentieth century this led to the establishment of sound archives, the earliest of which were located in New York City, Vienna and Berlin. Soon after, the first Russian collections were made, which later found their way to the sound archives of the Russian Academy of Sciences in the Pushkinsky Dom (Schiff and Marchenko, 1996). These sound archives contain about 10.000 wax cylinders of the

Edison phonograph and more than 500 old wax discs. In addition, an extensive fund of gramophone records exists and one of the largest collections of tape recordings of Russian folklore. These represent the history of Russian ethnography and contain a wide range of materials.

In the past, expeditions, often financed by philanthropists, were sent to various parts of the world to gather data, which were then returned to the archives for preservation and study. In 1897, for instance, on the initiative of the famous anthropologist and linguist Franz Boas, the Jesup Expedition set out from the American Museum of Natural History to examine evidence of similarities among the peoples of Siberia and the Northwest Coast of America. The ethnologists Waldemar Bogoras and Waldemar Johkelson made recordings of this expedition in Siberia (Jochelson, 1926). Many of these recordings are stored in the archives of the Pushkinsky Dom and they form one of the basic collections used in our INTAS projects.

4. Results of the INTAS projects on the languages of Russia

We were able to restore part of the many recordings in the Pushkinsky Dom and to make them available for further research, which is not only important for historical and cultural reasons, but also for the direct possible evidence of language change.

In our first INTAS project (1995 - 1998) we completed the reconstruction of sound archive material with the Zhirmunsky collection (Svetozarova, 1996). Zhirmunsky was a famous linguist who worked in Saint-Petersburg/Leningrad earlier in the twentieth century. One of his main interests was the study of German dialects spoken on the territory of Russia. In the period between 1927 and 1930 he and his assistants recorded many utterances on grammophone discs, in particular songs of German settlers. In the framework of this INTAS project, most of these discs have been copied on tape. Only a small part of the collection has been published so far and made available in a database and on CD ROM. This sound material, together with written data from the archives on the work by Zhirmunsky is now available for further research and it will provide interesting information on the life and culture of German colonists in the Russian Empire.

The sound archives in Saint-Petersburg also contain important data on Yiddish, the language of the Jews in Eastern Europe, which at the beginning of the twentieth century was used by millions of speakers in the Russian Empire. In the archives we found an unpublished manuscript on *The Ballad in Jewish Folklore* by Sophia D. Magid, together with material on corresponding wax cylinders. The manuscript is dated 1938, which explains why at that time it could not be published. Together with specialists in Saint-Petersburg, we are preparing a further exploration of the acoustic data in the sound archive and an edition of the book. This takes place in the framework of a Russian-Dutch project with the title: *Voices from the Shtetl: the Past and Present of the Yiddish Language in Russia*, for which we have obtained financial support from the Netherlands Organisation for Scientific research NWO (De Graaf, 1998b) for the period 1998 - 2001. Another aspect of this project is the present day situation of the Yiddish language, as far as it is still spoken by elderly people of Jewish descent

in St.Petersburg. In the interviews we made with these informants we find many interesting phenomena related to language change and code switching which can be described in the framework of modern theories on language contact.

Important activities related to linguistic databases in Saint-Petersburg also concern the many recordings of Russian dialects and minority languages in the Russian Federation, such as Nenets, Komi, Karelian, Vepsian, Khanti, Mansi, Tungus, Nivkh and others. One of our aims is the construction of a phonetic database of the languages of Russia which shall have many scientific, cultural and technical applications. In the beginning of 1998 we started a second INTAS project with the title *Sound Archives on the World Wide Web with Sound Recordings from Saint-Petersburg Collections*. We are preparing a database, which includes texts and sound material made in the first part of the twentieth century in the Arkhangelsk region of Russia. Of special interest for the project are the recordings of the Northern Russian dialects and the Komi and Nenets language made in that region. This database has been made accessible on the World Wide Web by implementing a dedicated Web-server at the Department of Phonetics of Saint-Petersburg University. In this way scholars around the world (linguists, phoneticians, ethnologists, musicologists and folklorists) can get access to this material (De Graaf, 1998c) by looking at the Internet address www.speech.nw.ru/phonetics/homepage.html.

The Northwestern regions of the Russian Federation are interesting from a linguistic point of view because of the fact that during the last centuries the original Uralic tribes living in this part of Europe have been assimilated by the Russian population. Because of this, there remain only small areas in which there are still speakers of Uralic languages such as Komi and Nenets. In a PhD project on *Language Contacts in the Russian North*, which has been approved by the Netherlands Science Foundation NWO for the period 2000 - 2004 the research group of Phonetics and Ethnolinguistics of Groningen University is investigating language contact phenomena which can be found in Russian dialects in this area and which are based on recorded sound material and on a contrastive study of Russian and these Uralic languages. In the framework of the project *The Nenets Language on the Web* we create a special internet site on spoken Nenets, where this language is represented by three different dialects.

In May 2000 another INTAS project on *The construction of a full-text database on Balto-Finnic languages and Russian dialects in Northwest-Russia* has been initiated together with research groups at the universities of St.Petersburg and Petrozavodsk (Karelia). The vicinity of St.Petersburg and the southern and middle parts of Karelia are a location where languages such as Vepsian, Ingrian, Votic, Ingermanland-Finnish and Karelian and various types of archaic Russian dialects are spoken. They represent an interesting sociolinguistic picture due to their long-term mutual interactions. The main objectives of this project are (1) to reveal and analyse various types of written sources for disappearing Balto-Finnic languages and archaic Russian dialects in Russia's Northwestern area, (2) to seek and record unparalleled modern text samples for these languages and (3) to construct

dictionary and full-text databases on Balto-Finnic languages and archaic Northwestern Russian dialects.

These databases can be used for scientific purposes (the study of language variety in Russia and language contact) and for the development of methods for language teaching in a bicultural environment. It provides a good opportunity for investigating ethnic and cultural processes in the contact zone of ancient Slavonic and Balto-Finnic languages.

In our projects we intend to combine the data from old sound recordings with the results of modern fieldwork, in order to describe the languages and cultures of ethnic groups in Russia. This information can be used for the preparation of text books on certain languages, collections of folklore and ethnomusicology and for the study of language contact, language change and migration movements. As an example of such fieldwork, I should like to mention the language of the Siberian Mennonites, which has strong links with the Low Countries (De Graaf and Nieuweboer, 1994). In 1998 this project resulted in a dissertation on Plautdiitsch that gives a description of various aspects of the language such as its phoneme system, the relation with other (Germanic and Slavic) languages, spelling problems of Plautdiitsch and language contact phenomena (Nieuweboer, 1998). We consider it an important task to record this disappearing Siberian Mennonite language and to describe it as precisely as possible.

5. Further projects on endangered languages

The extinction of languages is a process that takes place nearly everywhere in the world. The rich variety of languages that must have existed in the past is diminishing rapidly. As estimated by various linguists, in the next 50 years many of the 6000 languages, which are at present spoken in the world, will disappear (Kraus, 1991). At the moment children no longer use 20 to 50% of these languages, which makes their survival very uncertain. In our 1996 INTAS report it was stated that quite a few languages on the territory of the Russian Federation are under threat of total extinction and that measures should be taken to put an end to this process of degradation and dying out of languages. Linguists and ethnologists should work together with representatives of endangered languages in order to find solutions to these problems.

During fieldwork expeditions to Sakhalin and Northern Yakutia and in our other projects we have studied similar processes of language shift and language death for the aboriginal populations of Russia (De Graaf, 1992, 1996). It is a very important task for the linguistic community to make a registration of the last speakers of the endangered languages in interviews with good sound recording equipment. The results of modern fieldwork and the reconstructed data from sound archives will provide important information for the preparation of language descriptions, grammars, dictionaries and edited collections of oral and written literature. These can be used to develop teaching methods, in particular for the younger members of certain ethnic groups who do not have sufficient knowledge of their native language.

In certain cases the revival of endangered languages and cultures will be possible and for this purpose the contribution of linguists and ethnologists is of great value.

This work could further be stimulated by setting up exchange programs of Russian specialists with colleagues in Western-Europe and elsewhere, which will include the organization of fieldwork expeditions, further restoration of sound recordings, the construction of acoustic databases and the preparation of scholarly publications like textbooks and dictionaries. Similar to the way we started our INTAS projects, we can work together on a larger scale in order to study the languages of the Russian Federation.

One of our new project proposals called *Voices from Tundra and Taiga* considers the study of the languages and Russian dialects of Northern Russia (De Graaf, 1999). Here we plan phonetic/linguistic research on disappearing minority languages and Russian (archaic) dialects in Siberia and in the Far East of the Russian Federation, in particular on the island of Sakhalin. In February 2002 the Netherlands Organisation for Scientific research NWO has approved this project, while the international firm Sakhalin Energy is sponsoring it as part of its cultural program for the island.

Spontaneous speech and the reading of prepared texts will be collected that can be valuable for (ethno)linguistic as well as for anthropological, folkloric and ethnomusicological analysis. For that purpose, texts will be (video)recorded and analysed as to the art of story telling and language use. Described texts will be published in scientific journals and books with audiovisual illustrations on CD-Rom, and on the Internet. The materials will thus become available for further analysis to researchers working in the field of phonetics, linguistics, anthropology, history, ethno-musicology, folklore, etc. and will provide a possibility for the exchange of information with other institutions all over the world.

The results of modern field work and the restored data from sound archives provide important information for the preparation of language descriptions, grammars, dictionaries and edited collections of oral and written literature. These can also be used to develop teaching methods, in particular for the younger members of certain ethnic groups, who do not have sufficient knowledge of their native language, and thus make them aware of their heritage.

Our projects will contribute to the documentation and to the preservation of Russia's cultural heritage. This global collaboration with the use of information technology will make it possible to learn about the cultures and languages of Russia and it can provide new methods of teaching these topics, not only in universities, but also in special locations for the minority peoples. In this way the Russian Federation can develop a basis for sustainable multi-ethnic co-existence between Russians and the many other ethnic groups in the country.

6. References

De Graaf, T. 1992. Small languages and small language communities: news, notes, and comments. *International Journal of the Sociology of Languages* 94, p. 185.
 De Graaf, T. and Nieuweboer, R., 1994. The Language of the West Siberian Mennonites//RASK. *Internationaltidsskrift for sprog og kommunikation*, p.47.

De Graaf, T., 1996. Language Minorities in the Sakha Republic (Yakutia). *Ethnological Study of the Herders in Siberia*. Nagoya City University, p.165.
 De Graaf, T., 1997. The Reconstruction of Acoustic Data and the Study of Language Minorities in Russia. *Language Minorities and Minority Languages. Proceedings of the 6th International Conference on Minority Languages, Gdansk, July 1996/* Red. B.Synak and T. Wicherkiewicz. Wydawnictwo Uniwersytetu Gdanskiego, Gdansk.
 De Graaf, T., 1998a. The Use of Acoustic Databases in the Study of Language Change. Final Report 1995-1998. Groningen/ Saint-Petersburg.
 De Graaf, T., 1998b. Voices from the Shtetl: The Past and Present of the Yiddish Language in Russia. Project Proposal concerning Russian-Dutch Research Cooperation. Groningen, August 1998.
 De Graaf, T., 1998c. Sound Archives on the World Wide Web with Sound Recordings from Saint-Petersburg Collections. INTAS Research Proposal and Work Programme 1999-2001. Saint-Petersburg/Groningen.
 De Graaf, T., 1999. Voices from Tundra and Taiga. Proposal for a Research Project in Phonetics/Linguistics related to the Peoples in Northern-Russia. Groningen.
 Jochelson, W., 1926. The Yukagir and the Yukaghirized Tungus. The Jesup North Pacific Expedition/ Red. Franz Boas. E.J. Brill, Leiden.
 Nieuweboer, R. *The Altai Dialect of Plautdiitsch (West-Siberian Mennonite Low German)*. Dissertation Groningen University, 1998.
 Krauss, M., 1992. The world's languages in crisis. *Language*, T.68, p.1.
 Schiff, V.P. and Marchenko, Y.I., 1996. The folklore of the peoples of Siberia and the Far East in the Sound Archives of the Institute for Russian Literature (Pushkinsky Dom). *Archives of the Languages of Russia*. Reports on the INTAS project nr.94-4758. Red. T. de Graaf and L.V. Bondarko. Saint-Petersburg.
 Svetozarova, N.D., 1996. Zhirmunsky's collection of German folk songs in the Sound Archives of the Pushkinsky Dom. *Archives of the Languages of Russia*. Reports on the INTAS project nr.94-4758. Red. T. de Graaf and L.V. Bondarko. Saint-Petersburg.
 Witsen, N., 1966. *Moscovische Reyse 1664 - 1665. Journaal en aentekeningen*. Uitgegeven door Th.J.G. Locher en P. de Buck. Martinus Nijhoff. 's-Gravenhage, 1966.
 Witsen, N., 1996. *Puteshestvie v Moskoviyu 1664 - 1665. Dnevnik. Pervod so starogollandskogo V.G.Trisman*. Sankt-Peterburg.