



A SOUNDING DICTIONARY OF AN ENDANGERED LANGUAGE

Olga Kazakevitch
(Moscow)

1. Introduction

The paper presents the project "Local dialects of the Northern Selkups: a contrastive description and a data base of sound files" being realized at the Research Computational Centre of Lomonosov University of Moscow with financial support of the Russian Foundation of Fundamental Studies, grant N 01-06-80363. The project presupposes recording and processing materials in four sub-dialects of the Northern Selkup dialect still functioning at present: Middle-Taz, Upper Taz, Baikha, and Upper Tolka. The core of the project is A Sounding dictionary of local dialects of the Northern Selkups. It is regarded not only as an instrument for linguistic research but also as a linguistic resource, which can be used as teaching material in mother tongue classes.

1. Northern Selkups and their language

Selkup is the only living representative of the Southern subgroup of the Samoyed branch of the Uralic language family, spoken by the Selkups, an indigenous minority of Siberia. Counting 3612 people (according to the census of 1989), Selkups are dispersed on a large area of Western Siberia between the rivers Ob and Yenisei (see Map 1). The northern, or Tas-Turukhan, group of the Selkups resides in the Yamalo-Nenets autonomous area and in the Krasnoyarsk territory (see Map 2). and speaks the Northern (Tas-Turukhan) dialect. The southern, or Narym, group resides in the Tomsk region and speaks several different dialects. Nowadays the northern group counting 1883 people outnumbers the southern group (1347). The transmission of the southern Selkup dialects from parents to children has stopped since some decades. Now these dialects are spoken only by a handful of elderly people. The Northern dialect is better preserved, though the number of its speakers is being steadily diminished. A series of sociolinguistic surveys carried out in the places of residence of the Northern Selkups in 1996-2000 [1] revealed that the Northern dialect is now spoken by no more than 600 people and there are very few children among them. Practically all Selkup speaking Northern Selkups are bilingual with Russian as their second or (sometimes) their first language. In the elder generations there are also trilingual persons with one of indigenous minority languages of the area (Ket, Evenki, or Nenets) as the third (or as the second - in the latter case Russian being the third) language. Though a writing system has been developed for Selkup and now Selkup is taught at primary school as a subject, outside school it functions solely as a spoken language.

Map 1. Selkups Areas of Residence



CONGRÉS MUNDIAL SOBRE POLÍTQUES LINGÜÍSTIQUES
CONGRÈS MONDIAL SUR LES POLITIQUES LINGUISTIQUES
CONGRESO MUNDIAL SOBRE POLITICAS LINGÜÍSTICAS
WORLD CONGRESS ON LANGUAGE POLICIES Barcelona, 16-20 d'abril de 2002

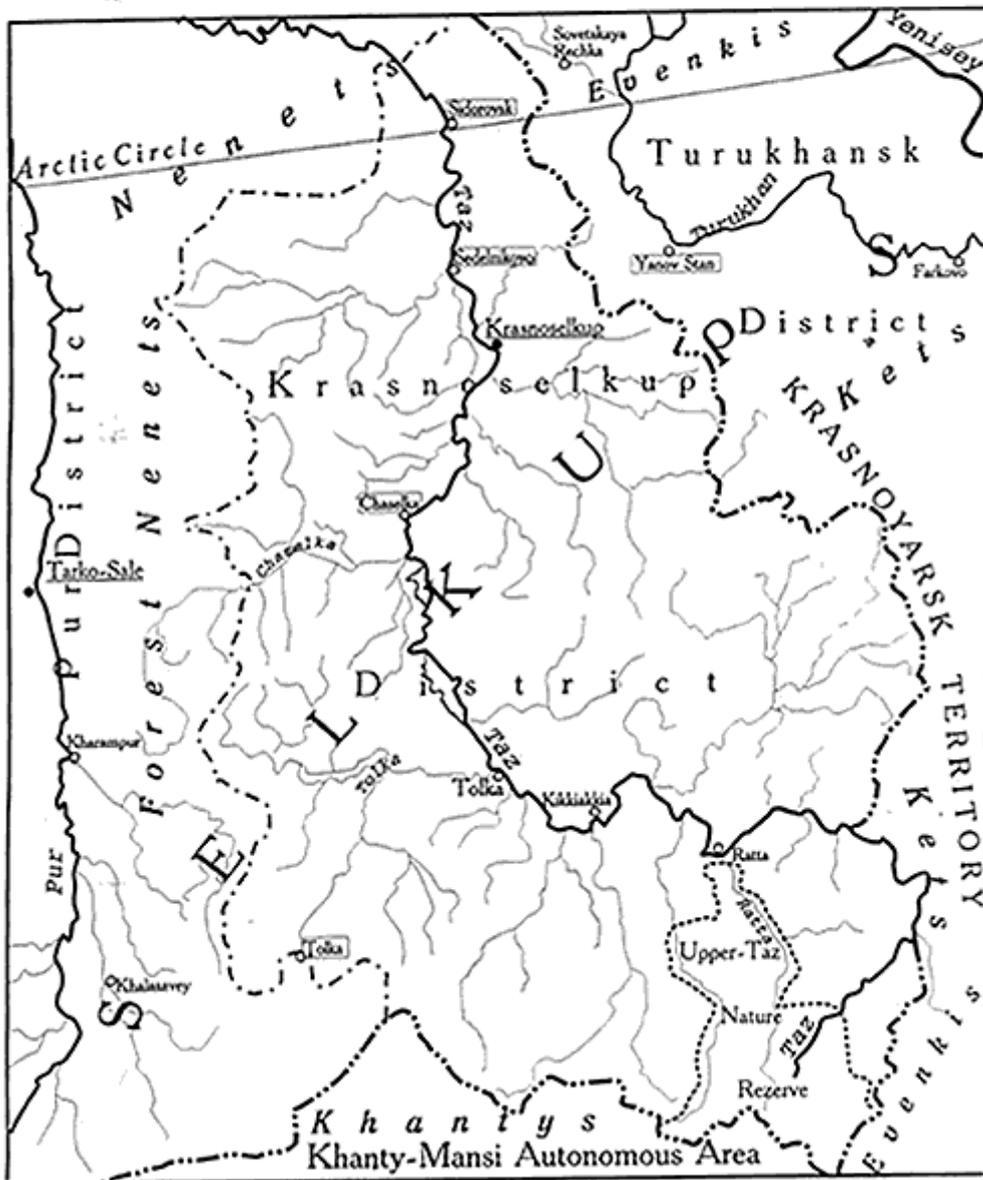


Map 2. Northern Selkups



Map 2. Northern Selkups

(Krasnoselkup and Pur Districts of the Yamalo-Nenets Autonomous Area and Turukhansk District of the Krasnoyarsk Territory)



The Northern dialect of Selkup is not uniform. It exists as a number of local sub-dialects. To-day at least four of them are still being spoken: the Middle-Taz sub-dialect (the village of Krasnoselkup, and the village Sovetskaya Rechka of the Turukhansk district - see Map 2), the Upper-Taz sub-dialect (the villages Ratta and Tolka of the Krasnoselkup district; some speakers of this sub-dialect live in the villages at the Yenisei), the Baikha sub-dialect (the village Farkovo of the Turukhansk district), and the Upper-Toplka sub-dialect (the village Tolka of the Pur district and the district centre Tarko-Sale).

2. Collecting materials for the Sounding dictionary

We present here the first stage of the project development: the Sounding dictionary of the Upper-Tolka sub-dialect. This sub-dialect was chosen to begin with because it had been much worse documented than the three other sub-dialects. The audio materials of the Upper-Tolka sub-dialect were being recorded in the summer 2001 during the expedition to the Pur district of the Yamalo-Nenets autonomous area in the village Pur Tolka and in the district centre Tarko-Sale by the author and her colleague researcher Irina Samarina from the Institute of Linguistics (Russian Academy of Sciences). A digital camera Sony DCR TRV 530E and a digital dictaphone Samsung SVR S1330 were used for this purpose.

Preparing for the expedition we compiled a Russian-Selkup wordlist sorted thematically, a sort of a thesaurus, which comprised about 2000 entries. As a basis we used the Russian-Selkup dictionary from (Kuznetsova et al. 1993) containing the Lexis of the Middle-Taz sub-dialect of the Northern dialect. The entries were distributed among the following 15 sections: 1. Body parts; 2. Nature; 3. Space; 4. Time; 5. Plants; 6. Animals; 7. People (terms of kinship, social roles, ethnic groups, spiritual culture); 8. Artifacts (house, clothes, transport, weapons, tools); 9. Food; 10. Abstract nouns; 11. Quantity; 12. Pronouns; 13. Mode of action; 14. Quality, characteristic; 15. Actions, processes, states. Every sections split into sub-groups.

During the work with a Selkup informant we gave him/her a Russian word from the list as a stimulus and we asked to pronounce a Selkup equivalent of this word. If the informant couldn't give a Selkup word we gave a second stimulus - a Selkup equivalent from the Middle-Taz sub-dialect. Quite often it helped, especially when we worked with an elderly person, a full speaker of his/her ethnic language but having not very high competence in Russian. As the sub-dialects of the Northern Selkup are inter-intelligible, as a rule, the informant recognised the Selkup word, commented on the wrong way the word had been pronounced, and then pronounced it in his own way, according to the norms of his own sub-dialect. We asked our informants to pronounce every word three times. Quite often before the recording we discussed with our informant some fragments of the word list so that he/she could recollect equivalents of the stimuli words (Russian and Selkup from a neighbour sub-dialect) in his native sub-dialect. It should be mentioned that during the interviews parallel with the recording of audio materials for the Sounding dictionary we tried to collect additional lexical and grammar material in the sub-dialect. That is why, as a rule, we did not satisfy with just getting equivalents in the local dialect variety pronounced but tried to document contexts with these words and to specify their meaning as well.

Here is an extract of the prepared word list (old information) and the answers received from one of the informants (new information).

Table 1. Fragment of the word list with new information

Old Information Russian word	Old Information Selkup equivalent (Middle-Taz sub-dialect)	New Information Selkup equivalent Upper-Tolka sub-dialect(Informant № 7)
волосы 'hair'	<i>opty</i>	<i>opty</i> <i>wyl'axyj opty</i> 'smooth hair', <i>pyrñaj opty</i> 'curly hare, <i>cumpy opty</i> 'long hair',

		<i>xo#mycy opty</i> 'short hair', <i>optäl' pütät(y)!</i> 'cut off your hair!'
<i>yxo</i> 'ear'	<i>ünkylsa</i>	<i>xo#ly</i> <i>ukkyr xo:ly</i> 'one your 2Sg ear' <i>xo#ly nom ammajimpa</i> '(your ear is frost-bitten)'

The word list was presented to eight informants and their pronunciation of the equivalents was audio recorded. The age of our speakers varied from 34 to 65, they had different educational background and different linguistic biographies. Three of them were over 50, and it's from these three informant that we audio recorded practically all the entries of the thesaurus. The younger speakers (under 50) couldn't recollect rather many Selkup words, so for them we have some lacunas in our education.

3. Dictionary Structure

Having input our audio materials into the PC and partly converting them we received files in wav format. The files were cut so that every separate file contained pronunciation of one separate word. The MS FrontPage 98 programs were used to create the document structure. The text information was prepared in MS Word 98. We used the font Lucida Sans Unicode to show the symbols of the IPA transcription. The demo version of the Sounding dictionary can be found on the site <http://www.infolex.ru>.

A dictionary entry can be entered through a Russian word, an English word, or a Selkup word (Middle Taz). The entrances do not sound. In the main zone of the dictionary the sounding files of the Selkup Upper Tolka words-equivalents of the entrance words, recorded from 8 speakers. (The Internet demo version of the Dictionary contains recordings from only four speakers). Every speaker pronounces each word three times. Every pronunciation is supplied with a transcription and, if necessary, grammar commentaries. The grammar commentaries become necessary if a speaker pronounces not an expected "dictionary" form (e.g. Nom Sg non-possessive for nouns, simple Infinitive for verbs) but some other form which seems more natural for him. Thus, our informants gave for nouns possessive and not non-possessive forms, and there could be forms of all the three persons Sg of the possessor. Sometimes our informants preferred to pronounce not a single word but a mini-context. In future we are going to include a zone of contexts into the dictionary structure.

At present the Dictionary includes materials of only one local sub-dialect of the Northern Selkup dialect. When we add materials of the three more sub-dialects it will become more informative but its structure will get much more complicated. Taking into account this issue we cut the present volume of the dictionary to 350 entries. We are going to check how it works with the materials of the four sub-dialects (we hope to have it in two years).

Table 2. Fragment of the inner structure of the Sounding dictionary of local dialects of the Northern Selkups

				Pur	Tolka	
Russian word	English word	Selkup word(Middle Taz)	Informant№ 1	Infrmant № 2	Informant № 3	Informant № 4
<i>yxo</i>	ear	<i>ünkylsa</i>	<i>qo</i>	<i>qo#-l'«</i> ear+Poss2Sg	<i>qo#-l'«</i> ear+Poss2Sg	<i>qo#-l'«</i> ear+Poss2Sg

челюсть	jaw	□#qyl«	□q-qyl l«	□q-qyl l«-ll«	□q-qyl l«-lly	□q-qyl l«-lly
			mouth+ Loc+bone	mouth Loc+ bone Poss2Sg	mouth Loc + bone Poss2Sg	mouth Loc + bonePoss2Sg

4. Conclusion

In the conclusion we'll dwell upon possible use of the Sounding dictionary. It is designed to be sounding archive of linguistic data of an endangered language (Selkup is just one of them). It can be used for computer analysis of the phonetics of the language. But it can also be used for practical purposes as teaching material in the ethnic language classes both on school and on the university level. Now it is already used at the philological faculty of Lomonosov University of Moscow within the course of Selkup. There is one more thing about the dictionary functions. The fact that sounding materials of an endangered language are present in the Internet seems to elevate the status of this language, and even if it does not enlarge the communicative function of the language it definitely influences its symbolic function and elevates its prestige which appears to be most important in the situations when the youth tend to abandon the language of their ancestors and shift to Russian.

References

Kazakevitch O.A. Minor aboriginal peoples of Russia: Language and ethnic self-identification // Proceedings of the international congress "Ethnicity and language community: an interdisciplinary and methodological comparison. Udine, 1998, pp. 307-323.

Kazakevitch O.A., Parfenova O.S. Yazykovaya i etnokul'turnaya situatsiya v Krasnosel'kupskom raione Yamalo-Nenetskogo avtonomnogo okruga [Language and ethno-cultural situation in the Krasnoselkup district of the Yamalo-Nenets autonomous area] // Yazyki Rossiyskoi Federatsii i novogo zarubezhya: status i funktsii [Languages of the Russian Federation and Ajaent States: Status and Functions]. Moscow, 2000, pp. 270-303.

Kazakevitch O.A. Yazyk verkhnetol'kinskikh sel'kupov: sovremennoye sostoyaniye i perspektivy sokhraneniya [the language of the Upper-Tolka Selkups: the present-day situation and perspectives of preservation] // Samodietsy. Materialy IV Sibirskogo simpoziuma "Kul'turnoye naslediyе narodov zapadnoi Sibiri" [Samoyeds. Papers of the IVth Siberian Symposium "Cultural heritage of the peoples of Western Siberia"]. Tobolsk-Omsk, 2001, pp. 270-273.

Kuznetsova A.I., Kazakevitch O.A., Ioffe L.Yu. Ocherki po sel'kupskomu yazyku. Tazovskij dialekt [Selkup. Taz Dialect]. Vol. 2. Moscow, 1993.

Notes

[1] Some results of these sueveys were published in (Kazakevitch 1998; Kazakevitch, Parfenova 2000, Kazakevitch 2000).