

**Interim Report (end of Pilot phase, beginning of mainphase)**

Disseminated March 2002

Title of Project: **Altai-Sayan Language and Ethnography Project  
(Tofa Project)**

Grant recipients: Endangered Language Fund  
Department of Linguistics  
Yale University  
P.O. Box 205977  
New Haven, CT 06520 USA

Responsible scientists: Prof. Dr. Bernard Comrie,  
Max Planck Institute for Evolutionary Anthropology  
Inselstrasse 22, D-04103  
Leipzig, Germany

Dr. K. David Harrison (principal investigator)  
Endangered Language Fund  
Department of Linguistics  
Yale University  
P.O. Box 205977  
New Haven, CT 06520 USA

Duration of Project: Pilot phase Aug. 2000 to March 2002  
Mainphase June 2002-August 2005

## **Abstract**

During our VolkswagenStiftung funded pilot project, we engaged in field documentation of the Tofa language and culture of southern Siberia. Presently spoken fluently by fewer than 30 people out of a community of 500, Tofa is rapidly vanishing. The Tofa community—living in isolation and poverty in a remote corner of Siberia—was generous and enthusiastic in their support of our documentation project. They hosted us at our field sites, provided logistical support, shared their language and traditions, and advised how we might contribute to the community's own goals and priorities for revitalization and preservation. In addition to the language, we undertook a comprehensive ethnographic study of Tofa culture, which is based on an animistic belief system and an economy of reindeer herding, hunting and gathering (Castrén 1857, Yevsenin 1919, Mel'nikova 1994). Lastly, we worked to document Tofa music, finding a rich tradition of unique song forms never previously documented.

In addition to Tofa, we worked to a lesser extent on two closely-related languages of the Altai-Sayan region: Todzhu (Republic of Tuva, under 200 speakers) and Tsengel Tuvan (Mongolia, under 1,000 speakers). These three languages are connected by genetic and areal affiliation, and by their connection to nomadic pastoralism and hunting. The endangerment status of Tofa, Tsengel and Todzhu in particular is, we feel, usefully viewed in the context of the demise of the traditional reindeer-herding and nomadic ecology of these peoples. Our finding is that Altai-Sayan languages and traditional ways of nomadic living are disappearing simultaneously; both are in urgent need of documentation.

The core of our project has been the scientific documentation of the language. We have written the bulk of a grammar of Tofa, which will be submitted for publication by Mouton De Gruyter in 2003. The grammar will draw on a rich collection of texts we collected in the field, as well as elicitation sessions with 18 native speakers. The texts represent a variety of genres: traditional stories, songs, personal biography, ethnobotany, hunting lore, speech directed to young children and spontaneous conversation.

We produced a large corpus of digitized video recordings (approximately 200 sessions). We have also posted samples of audio/video annotated data on our project website: <http://sapir.ling.yale.edu/~ASLEP/ASLEP.htm>. To conclude, we present a summary of **project results for the period** August 2000 through August 2001

- Three field expedition to Tofalaria,
- One field expedition to Mongolia,
- Five field expeditions to the Todzhu region
- Tofa grammar web version
- Tofa grammar print version (about 50% completed)
- Tofa children's book
- Filmed 40 hours of video tape
- Annotated over 60 linguistic sessions
- Created 7,000 word web lexica

## I. Publications / Description of the Project and its Results

### A. Publications and Dissemination of data:

We have presented the results of our research on Tofa and related languages at conferences and symposia. These include the LSA, WCCFL, The American Anthropological Society (planned), The American Oriental Society, and a symposium sponsored by the Finnish Linguistics Society. A partial list of these talks is as follows:

- Anderson, Gregory, Language Contact and Macro-Areal Typological Change: Complex Sentence Formation in Native Siberia. Paper presented at the 23<sup>rd</sup> Meeting of the Deutschen Gesellschaft für Sprachwissenschaft. Leipzig, February, 2001
- \_\_\_\_\_ Auxiliary Verb Constructions in Turkic (especially Tofa). Paper presented at the 211<sup>th</sup> Meeting of the American Oriental Society. Toronto, March 2001.
- \_\_\_\_\_ Switch Reference in Tofa: A Newly Discovered Vanishing Category. Presented at the Finnish Linguistics Society Symposium on Endangered Languages Helsinki, August 31-September 2, 2001
- Harrison, K. David, Altaic vowel harmony: Agree or Align? Talk given at WCCFL, USC, Los Angeles
- \_\_\_\_\_ The Tofa Project: Documenting an Endangered Language. invited talk given at MPI-EVA, Leipzig, April 2001. (with Sven Grawunder & Brian Donahoe).
- \_\_\_\_\_ Documentation, Endangered Languages and Technologies. Presented at the Finnish Linguistics Society Symposium on Endangered Languages Helsinki, August 31-September 2, 2001

We also wrote or collaborated on the following technical reports:

- Adaptive Responses to Institutional Collapse Among the Reindeer Herders of Tyva.* by Brian Donahoe. Published on our website, November 2002.  
<http://sapir.ling.yale.edu/~ASLEP/Adaptive%20Responses.htm>
- Many songs one tune: a Tofa field report.* by K. David Harrison. Published on our web site in December 2000 and reprinted in *Ogmios*, the newsletter of the Foundation for Endangered Languages. Nicholas Ostler, Ed. (31 March 2001).
- Approaches to Morphosyntactic Annotation.* by Arienne M. Dwyer, Hans-Heinrich Lieb & Gregory D. S. Anderson. Presented at Conference on 'Digitizing Endangered Languages', Santa Barbara, CA June 21-24, 2001.

We submitted for publication the following papers that include new Tofa data:

- Anderson, Gregory. 2001. Cislocative and Translocative in Tofa. Submitted to the journal *Ural-Altaische Jahrbücher*, 5/01
- \_\_\_\_\_ Subject Version and Object Version in Tofa Auxiliary Verb Constructions: Submitted to the journal *Turkic Languages*, 5/01
- \_\_\_\_\_ Micro-Areal Language Contact and Macro-Areal Typological Change: Complex Sentence Structure in Native Siberia. Submitted to *Linguistic Typology* 5/01
- Harrison, K. David. 2001. Vowel Harmony. In Phillip Strazny, Ed. *Encyclopedia of Linguistics*. Chicago: Fitzroy Dearborn.
- \_\_\_\_\_ 2001. Patterns, pervasive patterns, and feature specification. (with Abigail Kaun). In Urszula Klenhenz, Ed. *Studies in Distinctive Feature Theory. vol. 5. Phonetics in Phonology*, Berlin: Mouton de Gruyter.

The following book and monographs are in preparation:

- Anderson, Gregory and K. David Harrison. *Tofa*. (in preparation, currently more than 75% complete. We plan to submit this to the Routledge series).

Anderson, Gregory. *Case Marking in the Altai Sayan*. To be submitted to Harrassowitz Verlag, Turcologica Series. (In preparation, 30-40% finished).

\_\_\_\_\_. *Auxiliary Verb Constructions in Altai-Sayan Turkic*. To be submitted to Harrassowitz Verlag, Turcologica Series. (In preparation, 80% finished).

Using digital video was a completely new experience for us, one made possible only by support from the Volkswagen Stiftung and technical assistance from the TIDEL group in Nijmegen. In addition to shooting video in the field, we worked to reduce the turn-around time between field collection and the production of useable formatted and annotated data. We began transcription of the video tapes immediately in the field, usually producing a phonemic transcription on the same day the recording was made. Annotating sessions while still in the field allowed us to re-elicited or discussed new or unusual forms with speakers. It also allowed us to put materials more quickly into the DoBeS archives in Nijmegen. Upon returning from field trips, we began work on the remaining parts of the annotation: a morpho-syntactic tier, an English free translation, and a time code linking these to the audio/video signal. To date we have annotated nearly 40 sessions.

We also collected a large amount of Tofa lexical data, which we added to two online lexica we had previously compiled. The combined size of the lexica is about 11,000 words, with recently added Tofa words making up about 20% at the moment. The database will serve as the primary source for lexical comparison of these closely related languages, and also as the starting point for a massive comparative lexical database of all Altai-Sayan Turkic languages. This is a long-term goal of the Altai-Sayan Language and Ethnography Project, and one we expect to continue into the main phase.

In accord with the VW-Stiftung requirement for multidisciplinary research, we put together in the pilot phase a diverse field team: two linguists, one anthropologist, and a speech scientist/musicologist. We undertook an in-depth study of the music and ethnography of the Altai-Sayan region, designed to complement and inform our primary task of documenting the languages. The Tofa musical tradition afforded us with some of our most dramatic discoveries to date. Even closer to extinction than the Tofa language, Tofa music turned out to be quite unexpectedly different from any other song style documented in southern Siberia. We also found and recorded many examples of *sound mimesis* (hunting calls, animal imitations, etc.). Our results in documenting music are discussed in detail on page 8.

Our ethnographic study of the Tofa focuses on their traditional livelihood as reindeer herders (Vitebsky 1992), hunters and gatherers. We conducted a detailed sociolinguistic and ethnographic survey of 150 Tofa people to assess language use, ethnic identity, clan affiliation, and degree of involvement in traditional hunting/gathering activities. We are especially interested in those elements that link the language to the pastoralist (reindeer herding) ecology. We discovered, for example, that the traditional Tofa calendar (remembered now by only two elderly people) names months according to hunting/gathering activities associated with them: June was named ‘good birch collecting month’ and

November ‘sable-hunting month’. Documenting the Tofa language and investigating traditional reindeer ecology are thus highly complementary pursuits.

Beyond documentation, we took steps to support the indigenous communities in their efforts to record and preserve their language and culture. We provided financial and technical support for the Tofa Ethno-Cultural Center located in Alygdzher, the main Tofa village. This support included a cash grant, books, documentary video and photographs, and—at the specific request of the community—a short-wave radio and a photocopier. We produced a storybook of Tofa tales illustrated with drawings by local children (copy attached). Lastly, we delivered books to the Tuha and Tsengel communities in Mongolia that, although literate, had never or only rarely seen books in their native language.

### **I. Project team and field sites**

#### Project team:

Bernard Comrie, Ph.D.	MPI-EVA, Leipzig (co-applicant)
K. David Harrison, Ph.D.	Yale University (co-applicant)
Gregory D. S. Anderson, Ph.D.	University of Manchester
Sven Grawunder, Ph.D. candidate	Martin Luther University, Halle-Wittenburg
Brian Donahoe, Ph.D. candidate*	Indiana University, Bloomington
	*[Pilot phase only, not engaged in mainphase]

Our project was supervised and guided by Prof. Dr. Bernard Comrie at MPI-Leipzig. Prof. Comrie provided a host institution for us in Europe and advised us in planning and carrying out our project. The four principal researchers, listed above, were engaged in all aspects of the project, from field expeditions, to filming, to producing annotation and scholarly publications. David Harrison served as the principal investigator and was responsible for day-to-day operations and planning field expeditions.

#### Field sites:

Prior to the start of our project, we had not done any significant field documentation of Tofa, Todzhu, or Tsengel. We devoted considerable time in the pilot year to fieldwork. In 2000-2001, Tofa project members, individually and collectively, undertook a dozen field expeditions ranging from ten days to six weeks in duration. We worked at nine different field sites, six of which we had never previously visited. Most sites were at remote villages or nomadic encampments, where our speakers live. We also worked in the capital cities of Tuva (Kyzyl) and of Mongolia (Ulan Baatar) with native speakers and collaborating local scholars.



Map of field sites we visited during 2000-2001

<u>Region</u>	<u>Name of field site</u>	<u>language</u>
Tofalaria:	1. Alygdzher	Tofa
	2. Nerkha	Tofa
	3. Nizhneudinsk	Tofa
Tuva:	4. Ij	Todzhu
	5. Toora-Khem	Todzhu
	6. Saryg-Sep	Todzhu
	7. Kyzyl	Todzhu, Tuha
Mongolia:	8. Tsengel sumin	Tsengel
	9. Ulan Baator	Tuha

Pilot Phase: Time spent at field sites during 2000-2001 are shaded gray.

	2000					2001							
	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.
Anderson													
Donahoe													
Grawunder													
Harrison													

## **II. Scientific Results**

We summarize briefly our most significant scientific findings from our pilot-phase investigation of Tofa language, music, and ethnography.

### **Language**

During our field trips to Tofalaria, a number of surprising new issues came to light. These may be divided into: (i) 'new' phenomena previously unknown or only poorly known from published sources on Tofa, (ii) phenomena attributable to interference from Russian and, (iii) phenomena related to language obsolescence.

To the first category belongs the highly abstract system of vowel harmony. In Tofa, back vowels adjacent to [y] become fully front in their acoustics and articulation. But they remain back for purposes of palatal (backness) harmony, even when occurring in initial syllables of words. Similarly, we found speakers for whom front rounded vowels had become fully back, but who still classified these vowels as ‘front’ for purposes of harmony. This level of abstractness in palatal harmony has not been noted in any other vowel harmony language of the region.

Another significant ‘new’ phenomenon recently come to light in Tofa is a complex system of switch reference (Bergel’son & Kibrik 1987). In narrative styles, a verb in a complex sentence is marked for whether it has the same or a different subject as the following finite verb. We also documented modal particles *hoy* (expressing surprise and regret about an action), and *istig* (marking a necessary action), an adjectival suffix meaning ‘to smell of something’, and a verb-forming suffix with the same meaning. Finally, we documented the distribution and acoustic qualities of distinctive low pitch (Verner 1972, Bicheldei 1980), which Tofa employs as word-level pitch accent on a small class of words. The specifics and extent of these elements were previously unknown.

In the second category—phenomena not typical of Siberian Turkic and attributable to interference from the socially dominant Russian language—we documented the following:

- use of dative case to mark logical subjects in the necessative mood
- use of locative case to mark possession
- use of plural suffixes after higher numbers
- use of verbs in non-final position in clauses
- use of ‘when’ in clause-initial position in temporally subordinate clauses rather than a case-marked participle on the subordinate clause.

In the third category of recently discovered phenomena of Tofa—those relating to language obsolescence—we documented:

- vowel shifts and vowel mergers (loss of two front rounded vowels)
- reanalysis of vowel harmony by semi-speakers (phonetically back vowels count as front)
- replacement of less common modal forms with the future and the imperative
- collapses in the system of past and present tense marking
- the collapse of auxiliary verb constructions into a single one using the highly common auxiliary verb *ber/ver* ‘give’
- the collapse of certain uncommon complex sentence formants and markers used to denote switch reference
- collapse of kinship terms

In addition, there are other phenomena that recently came to light, the cause of which is not always easy to pinpoint. For example, despite the small number of fluent speakers remaining, we recorded significant variation in the pronunciation of certain phonemes across the speech community and within individual speakers. This may be an archaic feature of earlier dialect divisions, or may be the result of gestural shift due to language obsolescence. We plan future research to help shed some light on the nature of this variation.

## Music

When we began our field work, we did not know if anything of Tofa music survived. Given the small size of the speech community, we did not expect to find much. Of the Tofa song tradition, only a few texts written down in the 1970s (Stoyanov 1980), had been noted in the scientific literature. To our surprise, we found *nine* elderly Tofa people who sang a total of *three* distinct song styles:

- i. ‘traditional’ (a single melody, with special glottalization and register effects)
- ii. ‘shamanic’ (reportedly used in the past by shamans for healing rituals)
- iii. ‘modern’ (melody and texts copied from Russian songs)

Traditional Tofa songs rigidly follow a single tune (motif), varying only in individual *coloratura* applied by the singer. The songs are sung with special vocal modulations (i.e., a sharp, glottal interruption of every second syllable nucleus which introduces a voice register shift). We emphasize that this is a highly idiosyncratic style, and most likely a very archaic one native to Tofalaria. It does not resemble any of the known song traditions of neighboring peoples of South Siberia, and bears only a distant similarity to indigenous song styles of the far North (Kasten & Dürr 1999). The glottal modulations and register shifts render the words incomprehensible to the listener, as in a kind of speech disguise game. Further, the songs seem to employ archaic, intentionally obscure texts. Even when singers were able to report to us verbatim the words of a song they had just sung, they were often unable to interpret or parse the individual words, though they could sometimes paraphrase the general meaning of the text. We found ‘traditional’ style songs to differ among different villages. Each village employs a single variation on the same basic tune, but with numerous possible texts.

The ‘shamanic’ song style was performed by the two oldest singers in two different villages (V. Adamova, born 1914 and A. Mukhaeva, born 1914), but was unknown to everyone else. The two elderly women were able to sing a distinct song (with a different melody and singing style) which they reported as having been used by shamans for healing rituals. As no Tofa shamans survive, we cannot verify this claim. But, the style is sufficiently unique and remarkable to make a connection shamanic ritual plausible. If this is the case, this song would be the sole remnant of Tofa shamanic practice remaining in any form. A third, ‘modern’ and ethnographically less interesting style we recorded consists in Tofa lyrics set to Russian folk tunes.

Besides song, we documented native Tofa techniques of sound mimesis. We recorded younger speakers demonstrating the use of birch-bark horns and whistles and producing a variety of specialized hunting calls and vocalizations. Hunting calls follow rigidly defined motifs (which differ for different animals) but show individual stylistic variation. Older people also knew reindeer and horse domestication calls. These calls and whistles exemplify sound mimesis, an adaptive behavior by which indigenous people of the Altai-Sayan interact with the ambient acoustic environment by producing stylized imitative sounds designed to affect the behavior of wild and domestic animals (Levin 1990, 1999). Sound mimesis



among the Tofa represents not only an art form linked to animistic practices and beliefs, but may also be viewed as an archaic technology for managing natural resources.

### Ethnography

As part of our study of Tofa language and culture, we also explored traditional hunting and reindeer herding practices (Ingold 1980, Vitebsky 1992). Many of our linguistic sessions include discussions of ethnographic topics such as animal classification systems, hunting lore, uses of medicinal plants, weather phenomena and animistic beliefs (e.g. the bear cult). To complement this anecdotal ethnographic information, we conducted a systematic anthropological and sociolinguistic survey of 170 members of the community, Survey respondents answered questions about their ethnic self-identification, membership in kinship groups, language use, attitudes towards language shift, and degree of participation in traditional activities such as hunting and gathering.

### Video Sessions:

In the course of our field expeditions, we interviewed about 140 subjects (40 Tofa, 65 Todzhu, 25 Tsengel and 1 Tuha). From these we selected the more fluent speakers and made digital video recordings of approximately 45 subjects. From about 70 hours of raw footage, we identified about 160 distinct 'sessions' ranging in size from 30 seconds (e.g. a single song) to an hour (an extended interview). The sessions provide examples of many genres: spontaneous speech, structured interviews, songs, storytelling, ritual speech, humour, speech directed to children, genealogy, elicited sentences, sound mimesis, code-mixed speech, etc. We also produced about 40 non-linguistic, ethnographic sessions showing people engaged in traditional activities (shoeing a horse, playing a shaman's drum, saddling a reindeer, etc.)

We submitted videotaped sessions to the TIDEL team at Nijmegen for digitization, and the resulting digitized files were returned to us on DVD-ROM. Sessions were also placed on the ftp server for downloading and sharing among project members. Approximately 30 sessions have been annotated with the requisite three-tiers, while another 30 are partially annotated. Of the 60 annotated sessions, we have completed the time-linking of text to video signal for about 20 sessions. Time-linking has proven to be quite labor intensive using the MediaTagger software. The new EUDICO software (designed by the TIDEL group), which we are currently beta-testing, is expected to speed up time-linking significantly.

### Metadata:

In collaboration with our DoBeS colleagues, we adopted a metadata template designed to meet specific project needs and to conform to emerging standards for metadata such as IMDI (ISLE Metadata Initiative). We coded sessions according to language, genre, participants, content, etc. We have produced 242 metadata files to date, not only for video sessions, but also for ethnographic photographs, drawings, maps, etc.

Annotation tiers

DoBeS group members agreed that annotation of sessions will consist of at minimum two tiers: (i.) rendered text and (ii.) free translation. For our project we adopt a 3-tier model for most of our linguistic sessions (see image below): (i.) rendered text, (ii.) morphosyntactic tagging, and (iii.) free [English] translation. We are also currently developing additional tiers to annotate musical motifs of songs and special vocal features associated with singing styles (see page 12).

Phase one: Linking video to text:

During the pilot phase, we used MediaTagger™ software for time-linking annotation to digitized acoustic/video signal. We annotated about 30 sessions using MediaTagger, while another 30 are in progress. A sample frame of a session we annotated in MediaTagger is shown below.



Pictured above: T. Nurgaaz, a nomadic herder, speaking to David Harrison.  
© 2001 by the Tsengel Tuvan community and ASLEP.

The EUDICO tool

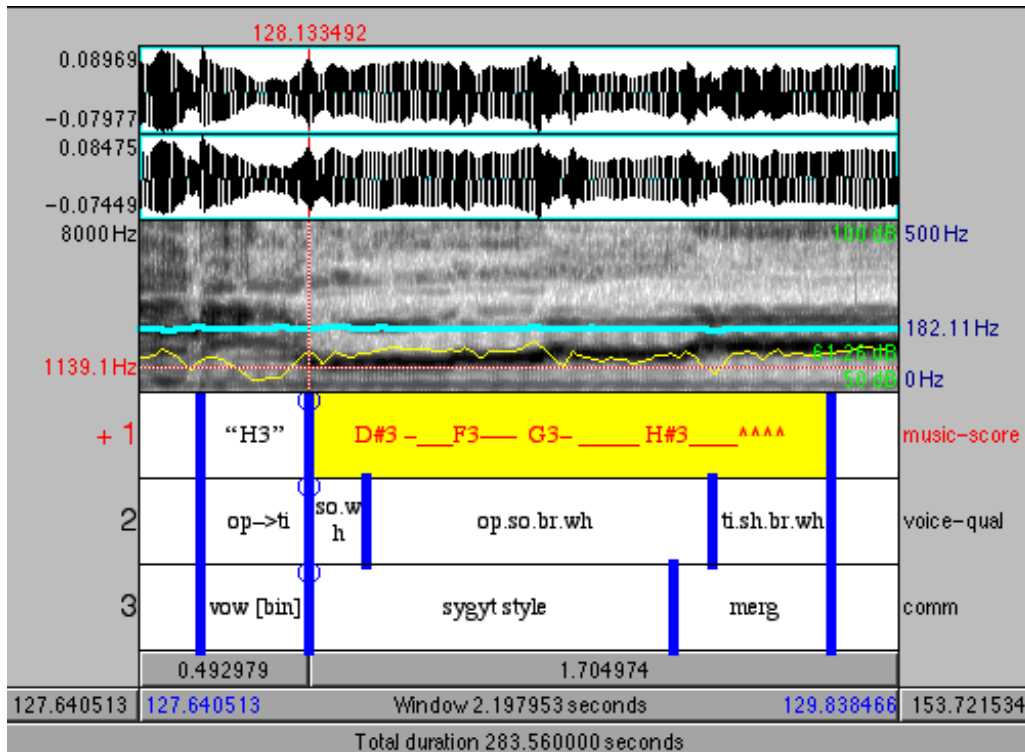
The TIDEL team has designed an audio/video annotation tool called EUDICO that should become a standard in the field. Currently our MediaTagger files are being transferred to EUDICO, and in the main phase we expect to do annotation and time linking exclusively with the new tool. The tool is currently in the beta-testing phase. Tofa sessions which we contributed to the archive in December 2000 were included in the preview version of EUDICO released June 2001. The sample frame shown below includes the video image, acoustic waveform, and three tiers of linguistic annotation.

The screenshot shows a video player window titled "TUDHDVN16Nov00-01CT-55". The window has a menu bar with "File", "View", "Player", and "Selection". The main video area shows a woman with dark hair, wearing a patterned shawl, looking down. Below the video is an acoustic waveform with time markers at 01:52.000, 01:53.000, 01:54.000, and 01:55.1. Below the waveform are three annotation tiers. The top tier is labeled "XXX-RT" and contains the text "Ūšpŭl turakēje šeni čokka siyirip čitar bolyan" with a second line "hazel.grouse day.night end NEG-DAT whistle-GER". The middle tier is labeled "XXX-TR" and contains the text "'morning and night the hazel-grouse whistled". The bottom tier is empty. The time markers for the tiers are 01:52, 00:01:53, 00:01:54, and 00:01.

In the top window a video image of Tatiana Nikolaevna telling a Tofa story in November 2000. The middle window contains the acoustic waveform. The lower window shows three annotation tiers synchronized to the video and audio signals. Note: some characters are incorrectly rendered here.

Advanced annotation

Taking advantage of EUDICO’s expandable tier structure, we are exploring ways of annotating ethnographic sessions that include visually or acoustically expressed content other than language. We require additional notation to describe the highly specialized song forms such as ‘throat singing’ and ‘overtone singing’ (Levin & Edgerton 1999). These are unique to the Altai-Sayan region of Siberia, e.g. the Tuvan, Xakas, Altai and Tsengel peoples, and are only superficially similar to other ‘throat’ singing styles such as that employed by Tibetan monks. Sven Grawunder is developing a speech event tier that will encode vocal features and techniques: pitch, register, timbre, overtone singing, etc. (Grawunder, in preparation). He will adapt a musical notation tier to annotate song motifs. Sven records specialized vocal techniques in the field using three microphones (one for the voice and one for the chest resonances) and an EGG laryngeograph on the outside surface of the throat. Using Praat™ software he can link the laryngeograph signal to the vocal acoustic signal for a full annotation of specialized singing techniques. A sample segment of Tuvan overtone singing annotated in Praat is shown below:



The acoustical analyses window includes a narrow band sonogram, a F0 pitch and intensity analysis. Tier 1 (music-score) shows a numeric score based on usually phonetic scales (Titze 1991). It is adapted to our signal so that indifferent notes are shown in quotation marks and horizontal lines indicate steps during the onset phase of a sung phrase or ‘tone’. The second tier describes the varying voice qualities based on a system of auditory/auditive and articulatory categories, including larynx position, glottal oscillation

quality, overtone qualities, voice register, etc. (Gall & Berg 1998, Grawunder 1999). The ‘comm[ent]’ tier includes general notes of a stylistic typology (e.g. ‘vow[el]’, ‘sygyt’) and also remarks on events in the acoustic analysis window (e.g. ‘merg[e]’). These last two issues could also be extended to two new tiers plus additional tiers to phonetically transcribe and gloss the sung text. These specialized annotations prepared by us in Praat will be imported into EUDICO by the TIDEL team.

Ethnographic content:

In addition to linguistic content, we recorded many sessions rich in ethnographic content. Sessions include demonstrations (often accompanied by speech) of how to build a yurt, how to shoe a horse, how to play a birch-bark hunting horn, how to perform a shamanic seance, etc. We also collected examples of storytelling, singing, dance performance, joke-telling and classroom teaching. In the case of Tofa, we focused on traditional hunting and reindeer herding activities.

Tofa speaker Albert Tulaev (born 1938) discusses traditional hunting techniques.



### Online lexica

We are currently in advanced stages of editing two online lexica:

1. **Tofa**-Tuvan-English-Russian                      about 6,000 entries
2. **Tofa**-Tuvan-Xakas-German                      about 3,000 entries

These lexica are housed online at the WebCelex online lexicon site which the TIDEL team is adapting to meet the specifications of various DoBeS projects. Among the lexica designed by DoBeS projects, our lexica have perhaps the simplest structure. The bulk of our database consists in Tuvan lexemes (collected prior to the start of our project) with English and Russian translations. With the help of the TIDEL team, we successfully imported these legacy materials into the WebCelex database. We have now begun adding Tofa and Todzhu lexemes into the lexical database. Our native speaker consultant in Tuva, Mr. Kyzyl Simchit, assists us in editing lexical materials online. In the main phase we will also add Tsengel, Tuha and Altai lexemes. The juxtaposition of these closely related languages in a single database will allow for comparative analysis. We envision them as an easy-to-use, searchable resource for languages which up to now have had no lexical materials or materials published only in Russian.

In the main phase we will add another layer of structure to our currently very simple lexicon. Words will be additionally coded by semantic domain (e.g. 'theme'). Semantic domains we are coding for Tofa include hunting, the bear taboo, body parts, musical instruments, etc. Users will have access to a master list of all the semantic fields and can by thus call up entire semantic sets.

### Back-up archiving and Data sharing

Archiving by the TIDEL group notwithstanding, we assume responsibility for long-term archiving of the data. We have made arrangements with the Yale University Department of Linguistics to permanently archive our digital media on DVD-ROM. In addition, we have agreed to collaborate with the NSF-funded E-MELD project by contributing a corpus of annotated video sessions to the E-MELD online archive to be housed at The Linguist List. Within E-MELD, our sessions will serve as examples of a text plus video annotation format. We also hope to take advantage of data storage possibilities offered us by the Linguistic Data Consortium. Finally, we are considering ways of sharing some of our field data with a research institute in the Republic of Tuva.

## **IV. Community assistance**

### The Tofa Ethno-Cultural Center

One goal of our pilot project was to provide direct (both financial and material) assistance to the communities where we conduct field work. We earmarked about € 3.000 of our pilot budget for this purpose. About one third of this was paid directly to linguistic consultants in the field. We solicited advice from Tofa community leaders their on to allocate the remaining amount. At their recommendation, we

agreed to channel most of these resources to support the **Tofa Ethno-Cultural Center** in Alygdzher village. We chose to do this rather than establishing a separate ‘Tofa language center’ as described in our pilot proposal.

The current center has been in existence for about five years, but receives no financial support from any source. Up to now, it operated on a zero budget, staffed by volunteers. Despite limited resources, the center organized a number of important cultural activities in the village: there is a youth dance troupe, a birch basket-making workshop, and an annual cultural festival celebrating traditional song and games. The center’s two-person staff (consisting of one Tofa and one Russian who is deeply involved in Tofa culture) collects and displays cultural artifacts, ethnographic photos, and books written in or about Tofa. They also organize regular trips to the forest to collect medicinal plants. Finally, they provide logistical and intellectual support for our field visits.

In June 2001 we delivered € 1.800 in cash to support activities of the Tofa Ethno-Cultural Center. With the support of our project, the center has already begun to fulfill the functions we envisioned for a language center. It serves as a repository from which books, video and audio materials we have contributed circulate in the community. It provides a forum for video screenings, social gatherings, storytelling and informal meetings. We use it as a research space and a place to meet with our language consultants while we are in the village.

In sum, we have endowed the Tofa ethno-cultural center with its own budget, a repository of video and printed linguistics materials, and a collection of photographs. Part of the budget has been used for training and employing two Tofa community members who have an active interest in pedagogy and language preservation. We have also provided a catalyst for local people to gather at the center and continue to exchange ideas about language preservation. In the main phase, we expect to supply the center with its own video camera and train local people to record cultural data on videotape. Working at the center, they will assist in data collection, transcription and archiving.

Our support for the center is a concrete gesture to the community, signaling that we are not there just to record and take away their cultural wealth, but to support and enrich their cultural life. In sponsoring the operation of center, we ensure that our work makes a tangible contribution to Tofa cultural revitalization. We will continue and increase our support in the main phase (see budget) allowing the center to function as a community gathering place, an archive and a scientific field center.

#### Gifts and payments to language consultants

At the request of the community, in June 2001 we delivered a small photocopier machine, a short-wave radio for the use by the reindeer herders, and art supplies for the children. We also delivered Tofa books, videotapes of elderly Tofa speakers, and ethnographic photographs.

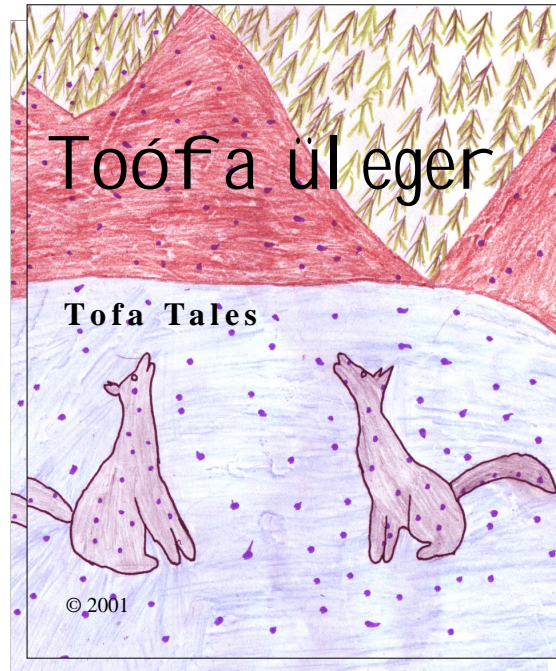
We also provided community assistance in the form of direct payments (cash, food, or small gifts) to our linguistic field consultants. In communities where the average monthly income is well below



€ 100, such payments have a dual benefit. First, they provide tangible economic resources to people who have shared with us their intellectual wealth. Second, they may serve to raise the relative prestige of the ancestral language in the eyes of the community.

### Pedagogy

We produced a children's book containing stories in Tofa with a Russian translation. This is the first bi-lingual book of Tofa stories and only the fourth Tofa book ever published. Anyone in the community can now read traditional Tofa stories in Russian, and, if they are interested, can view the original Tofa text as well. Tofa children drew the artwork for the book, based on their interpretation of traditional stories read to them in Russian. Ms. Tatiana Kangaraeva, the local Tofa language teacher, plans to use the book with her 8 to 10 year-old pupils. A copy of the book is included with this proposal.



© 2001 by Tofa community and ASLEP.  
Cover art by T. Tulaeva (born 1988).

### Language revitalization?

While we consider it unlikely that Tofa, now moribund, will be revived, we view it as part of our ethnical responsibility to enable the community's efforts at language preservation. In addition to tangible steps like publishing the storybook, we have gleaned some intangible results in this respect. Many elderly people have expressed to us their keen sense of regret that the next generation fails to value or cherish the language. Since commencing fieldwork in Tofalaria in November 2000, we have witnessed shifts in some younger Tofa people's attitudes towards learning and using the language. A 26-year old whom we employed to assist with the sociolinguistic survey began to study the language in earnest and within six months she was able to produce simple sentences. Several people in their late 30's or 40's who had at best a passive knowledge of the language began to share with us certain words or phrases they were able to remember. Their halting spoken Tofa turned out to be of significant theoretical interest to our study of cross-linguistic grammatical interference and obsolescence. We documented several new phonological and syntactic patterns in semi-speakers that we attribute to language interference. But for our active interest, these semi-speakers were unlikely to ever utter a word of Tofa. In collaborating with us, some were able to overcome their inhibition about speaking the language less than fluently.

We have seen that the presence of foreign researchers who place a high value on the language can increase its social valuation in the community. We hope to continue arousing younger people's interest in the language, perhaps slowing down the process of loss and thereby allowing more time for documentation.

Summary of Tofa community assistance during pilot phase:

Direct grant (€ 1.600) to Tofa Ethno-Cultural Center  
Linguistic materials (videotapes, texts)  
Tofa story book  
Photocopier and supplies  
Short-wave radio for the reindeer herders  
Material assistance (tools, clothing, medicine, etc.)  
Ethnographic materials (photographs, survey results)  
Cash and gifts to language consultants (approx. € 1.000)

Public relations activities

The funding principles state the following: “[t]he Stiftung welcomes the recipient of a grant publicizing a project and the Stiftung's decision to fund it.” In accord with this principle, and at the request of Dr. Vera Szöllösi-Brenig, we communicated with several journalists to discuss our own research program and the goals of the foundation in supporting endangered language work. We hope to help raise public awareness of language endangerment and the need for documentation. These contacts resulted in the following press articles that specifically mentioned the Stiftung and one or more of the DoBeS projects.

The Boston Globe

*Vanishing tongues: Scientists fight to save world's disappearing languages.* by Gareth Cook  
November 5, 2000

Süddeutsche Zeitung

*Wortinseln in Datenstrom.* by Harald Staun  
December 12, 2000

The Sydney Morning Herald

*A Desperate Race in Anyone's Language.* (reprint of Boston Globe article)  
December 20, 2000

Ogmios (Newsletter of the Foundation for Endangered Languages)

*Many songs, one tune: A Tofa field report.* by K. David Harrison  
March 2001.

GEO Wissen

*Rettung durch das WWW: Forscher wollen mithilfe des Internets vom Verschwinden bedrohte Sprachen und Dialekte bewahren*  
April 2001

Focus 17 June 2002  
[feature science article, in German] by Christian Weber  
photos by Thomas Hegenbart

With the encouragement of Dr. Szöllösi-Brenig, we agreed in June 2001 to take along to Tofalaria a science correspondent from Focus magazine and a professional photo-journalist. The presence of the journalists had, we feel, a positive effect on community morale. In accord with community wishes and at our request, the photographer agreed to provide copies of his photographs to the Tofa Ethno-Cultural Center, making a significant contribution to the photo archives of the center.

## **V. Community ownership and dissemination of data**

### Community ownership of data

In return for sharing their linguistic and cultural wealth with us, Tofa community elders have made specific requests regarding our use of the data. We share with them an understanding that all recorded data is intellectual property owned by the community and its individual members. As scientists, we are merely co-caretakers of this data. We obtain appropriate consent for archiving and dissemination of a person's voice and image, and we inform our consultants of the eventual location and uses of the recordings. We also pay our consultants for their time and expertise, typically in small amounts of cash or gifts. Finally, we agreed to fulfill community elders' requests (made in November 2000) to return the following to the community,

- copies of still photographs
- selected portions of videotaped material
- copies of printed texts (see attached Tofa storybook)
- compiled results of sociolinguistic surveys

During our June 2001 expedition, we returned video, audio, photographic and printed materials to the community. Some materials were deposited in the Tofa ethno-cultural center, as a part of our support for the center's activities. Other materials were given directly to individual consultants or their families.

While we were in the field in June 2001, an elderly Tofa speaker whom we had filmed in November of 2000 died. We were worried because we had brought along with us for distribution to the community a videotape which showed this person singing and talking. To our relief, we discovered that there was no death taboo that prevented people from viewing recently deceased persons on video tape. To the contrary, Tofa people who viewed the tapes were quite moved at the thought that they had images of a community matriarch preserved on video. Several community leaders emphasized their strong wishes to receive from us copies of our language documentation materials for posterity. We have agreed to continue doing so for the duration of our project.

## **Cooperations**

We were actively assisted in our field documentation by a group of young, indigenous scholars, including a native Todzhu, two Tuvans (Republic of Tuva), and two Tsengel speakers (Mongolia). We provided these colleagues with a laptop computer, the use of video and audio equipment, training in data processing (annotation of the audio/video signal), and training in telecommunication (internet use). We were particularly fortunate to have Mr. Alex Sharapai, a native Todzhu who, with our help, collected extensive lexical materials towards his Master's thesis. Ms. Tsetsegdar of Mongolia has collected a 2,000+ word lexicon of Tsengel and delivered books purchased by our project to that isolated community. Mr. Kyzyl-Maadyr Simchit, a native Tuvan philologist currently earning his Master's degree in Turkey, learned to do IPA transcription, interlinear glossing and annotation of video recordings. We provided him with a laptop PC for use in the field and receive his work via the internet. Ms. Polina Seren, a native Tuvan turcologist, worked with us on transcribing and annotating video recordings of speakers.

### Field-based collaborators:

Alexander Sharapai, M.A. candidate	Pedagogical Institute, Republic of Tuva
S. Tsetsegdar, M.A.	Pedagogical Institute, Republic of Tuva
Kyzyl-Maadyr Simchit, M.A. candidate	Atatürk University, Ankara, Turkey
Polina Seren, M.A.	Humanities Research Institute, Rep. of Tuva
G. Zolbayar, Ph.D.	Tsengel Cultural Institute, Mongolia

We also carried out collaborative work within the DoBeS group, as follows:

### Morpho-syntactic tagging working group

Significant collaborative work was carried out under the auspices of the DoBeS working group on morphosyntactic annotation. A number of theoretical and practical issues were discussed, and the results summarized in a presentation at a conference on 'Digitizing Endangered Languages' in Santa Barbara, California during late June of 2001. The paper 'Approaches to Morphosyntactic Annotation' was co-authored by Arienne Dwyer of the Salar/Monguor project, Hans-Heinrich Lieb of the Aweti project, and Gregory Anderson of the Tofa project.

### Ethnographic data working group

In discussions with DoBeS members, we explored ways of organizing and annotating non-linguistic, ethnographic data within the archive. We have amassed a rich collection of photographs, maps, and drawings. These shed light on the cultural complex in which the languages we investigate are spoken. We are working to devise appropriate methods to build an interactive web archive of visually appealing, informative materials while protecting the intellectual and cultural wealth of the community.

Main phase responsibilities of individual project members:

- Comrie: Prof. Bernard Comrie oversees the ASLEP project, especially the linguistic aspects of our research. He has oversight of all funds during the main phase. He hosts project members during research visits to MPI-EVA in Leipzig and acts as a liaison to the Volkswagen-Stiftung.
- Anderson: Greg specializes in Turkic linguistics (especially morpho-syntax and language shift). He is primary compiler of the main online lexicon and is co-author of the Tofa grammar. He serves on the DoBeS working group on morphosyntactic tagging. He annotates linguistic sessions and presents scientific results at symposia and in published papers.
- Donahoe: [Pilot phase only, not engaged in mainphase] Brian specializes in cultural anthropology and the ethnography of reindeer herding. He films and annotates linguistic and ethnographic sessions, creates metadata, and coordinates the work of our indigenous field assistants. He organizes community assistance projects, conducts socio-cultural field surveys and presents results of our work at anthropology conferences. Finally, Brian has produced detailed photographic documentation our field work and has given copies of this work to the local communities.
- Grawunder: Sven specializes in musicology and speech science (phonetics). He films, records and annotates song genres and linguistic sessions. He does acoustic analysis of sound mimesis and specialized vocal techniques (e.g. throat-singing, overtone singing) peculiar to the Altai-Sayan region. He is sole author of the Tuva-Tofa-Xakas-German online lexicon.
- Harrison: David specialized in linguistics (phonology) and ethnography and directs the day to day operations of the project. He is responsible for planning and executing field expeditions. He maintains contacts with the TIDEL team at Nijmegen and external contacts with the press, with host country institutions, and with collaborating scholars. He films and annotates linguistic and ethnographic data, and is co-authoring (with Greg Anderson) the Tofa grammar and the online lexicon. David is also the designer of the project web site and coordinates our backup data archiving at Yale.

**Information on further activities.**

In the main phase for which we have been awarded funding (April 2002-Summer 2005), we plan first and foremost to continue our work on documenting Tofa language and culture. We will visit the third Tofa village in 2002 and meet with all the speakers we have not yet met (we estimate their number to be about 12 to 15 persons). We will continue to provide significant community assistance in the form of direct grants, technology, support for community-based revitalization efforts, and pedagogical materials.

Also in the main phase, we plan to increase our investigation of the related small languages of the Altai-Sayan region. We have already collected, annotated and deposited in the DoBES database in Nijmegen material from each of these three languages (in addition to the Tofa materials). In the proposed main phase, we will build a comprehensive, accessible, and diverse database of these four critically endangered languages. For each of these speech communities, we will also continue to collect and annotate comparative ethnographic materials in the form of video, still photographs, music recordings, and surveys.