From the Field to the Net: Cataloguing and Digitising Cultural Research Material

*by Tiina Mahlamäki**

Introduction

My purpose in this paper is to describe the progression of research material, mainly tape-recorded interviews, from the field to the Sound Archive of Folklore and Comparative Religion i.e. the TKU Archive at the University of Turku, Finland. I will describe briefly the processes of cataloguing the interview material and the progression of the original analogue material to on-line use on the Internet. Also, issues in these key areas will be highlighted: the process of digitisation; the creation of databases; and associated with equipment, technology, and ethical questions.

The TKU Archive

The TKU Archive, founded in 1964, belongs to the Department of Cultural Studies at the University of Turku, Finland. The Department consists of four disciplines: archaeology, comparative religion, ethnology and folkloristics. They have vast collections of traditional material, mainly collected from different areas of Finland, but the fieldwork projects in Russia, Estonia, Peru, China, and India have also produced large collections of research material. (Mahlamäki 2000.)

In the mid-1960s new optimism concerning fieldwork became a significant part of folkloristic research in the Nordic countries. It was also a time of new fieldwork ideology. In Finland, the new anthropologically oriented folkloristics acquired a foothold especially in Turku. The focus shifted from folklore texts to social and psychological aspects of folklore in living communities. Many new questions arose. For instance, the problems of learning, performing, transmitting, and interpreting oral tradition became of interest to scholars, as well as discovering new genres of oral tradition. (TKU/A/00/79: 1; see Nyberg & al. 2000: 491-494.)

During the 1960s, the Section of Folkloristics and Comparative Religion started several research projects that were directed towards developing research methodology, fieldwork and archiving techniques. The disciplines are very fieldwork-oriented and the main projects have been Saami folklore; life-stories and oral history of Ingrians; Baltic-Finnic laments; ethnomedicine in Peruvian Amazon; oral epics in India; and religious, ethnic and cultural groups in Turku area. The TKU Archive is the third largest traditional archive in Finland, and its material consists mainly of approximately 11,000 hours of interviews and numerous photographs, slides, manuscripts, and other original and copied material. (Herranen & Saressalo 1978, 65; TKU/A/00/79:8; Mahlamäki 2000.)

From the Field to the Archive

From the beginning, the cataloguing system of the TKU Archive differed from that in other traditional archives because the recording tape was defined as the basic archive unit with its own accession code. Each archive unit was then divided, by the means of writing a recording protocol, into data units that were understood as analytical units separate from the entire tape. They were also considered to be the smallest coherent units to be referred to in research reports. Usually the data unit was the unit of content which was most often determined by a single folklore motif. (TKU/A/00/79: 3, 12; Mahlamäki 2000.)

For over two decades, each and every research project has, more or less, followed its own principles in archiving materials in the TKU Archive. By the end of the 1980s, the need arose to renew and standardise archiving practices.

In 1988, researchers in the Department edited a new kind of cataloguing system based on a filing card named COLLCARD (collection card). The COLLCARD filing system was first of all a medium for storing archive material on computer. Its content and structure were based on researchers' actual fieldwork experiences. COLLCARD is an excellent tool for use in the field. It is a notebook in which one writes the basic data needed for archiving an item. With COLLCARD it is possible to start the archiving process immediately during fieldwork. It also makes it possible to catalogue all kinds of material according to the same system. It contains the most important basic data that must be known about a folklore item of scientific value. COLLCARD was first a paper version that was used in the field and later was converted to a WordPerfect macro that was used when cataloguing the material into the TKU database, created in 1989. Now the COLLCARD form can be found and filled in on the Internet and can then be sent to the Archive by e-mail. (Huttunen 1992; Huttunen & al

1991; Rajamäki 1989, 35; Mahlamäki 2000.)

The COLLCARD filing system has proven its efficiency during the last decade. Firstly, when using COLLCARD during fieldwork, half the archiving process is done before returning home. Secondly, the uniformity in the note-taking technique helps combine the observations and data of other researchers on the same project even if the language or culture is unfamiliar to them. Thirdly, the COLLCARD filing system can be used in training new and inexperienced researchers. By using COLLCARD during fieldwork, the collector becomes practised in asking the informant the basic questions that might otherwise be forgotten or remain obscure. (Rajamäki 1989, 38-39.)

The Saami Folklore Project
I now describe in more detail the Saami folklore research project launched in 1965 in different parts of the Lapland of Finland, Sweden and Norway. The ideological background was to prove the existence of living Saami folklore (as opposed to the age-old literary sources), and save it in the Archive. The main target of the in-depth research project that lasted for many years was the village of Talvadas1 by the river of Teno. Three neighbouring villages were selected to serve as comparison villages. All the inhabitants over 16 years old in the village of Talvadas were inter-viewed several times over several years. (The Pilot Project 2000; Nyberg & al 2000.)

During the 1960s and 1970s, the project collected approxi-mately 1000 hours of recording tapes from the whole Saami area. Most of the sound tapes are in Saami language, but some of them are also in Finnish and in Norwegian. All the material has been archived in the TKU Archive. Additional fieldwork trips have also been made in the 1970s, 1990s, and at the beginning of the year 2001. Thus, a remarkable amount of new material has been added to the old material corpus. (The Pilot Project 2000; Mahlamäki & Enges 2001.)

In addition to the sound recordings, registers and transcrip-tions, the collection also includes photographs and slides along with a rich and varied collection of manuscripts and maps. During the fieldwork period all the material was catalogued manually. At the moment, all the photographs and slides and also over 200 sound recordings have been re-catalogued into the Talvadas-database (see below). The process of cataloguing into the database is still in progress. (The Pilot Project 2000.)

This collection is probably the most extensive sound collection of Saami folklore in existence. It forms a valuable and irreplicable whole that would be impossible to recreate today because the oldest informants, born in the 1880s, are no longer living and many dialects of the Saami language have now vanished. From the point of view of folkloristics, the material is a very rare 'thick corpus' that

makes it possible to study organic variation and various meanings of folklore in a local community. (The Pilot Project 2000; Mahlamäki & Enges 2001.)

From the archive to the Internet
The material of the Saami folklore project is still actively used at the University of Turku and there is a lot of interest in it and co-operation in connection with it from several universities in Scandinavia. But there are two kinds of problems concerning the optimal use of the material: access and preservation. The material is situated at the TKU Archive and all researchers interested in it have had to come to Turku. Also, the original material is suffering from both the unsuitable archiving conditions and use and copying of the material.

One solution for both the problems is digitisation of the research material. In a digitised form the material could be approached via an on-line connection all over the world and the original material would remain safe for use by future generations. (The Pilot Project 2000.)

Digitisation is, of course, not the only way to save the vanishing tapes. Just recording them to new analogue tapes would do the same. But digitisation will provide many advantages in addition to saving the information. Use of the material is not dependent on time and place, because the material is available on the Internet. It will also be faster and easier to use large amounts of research material because it can be handled, searched and studied by using the Archive database. Several researchers can also use the research material at the same time. And last but not least, the original material will be safe from damage caused by copying and usage. (Kurkela 1999; Saarinen 2000, 119-120.)

Talvadas-database
The new Talvadas-database, created in 2000 at the begin-ning of the digitisation project, is based on the TRIP Highway program. It is now possible to search for and retrieve the COLLCARDs and the material of the Saami folklore project via the Internet. Because of the legislation concerning personal data, and also for reasons of research ethics, access to the Talvadas-database is not freely available to all users. The database is in a closed net environment and access requires a special password that is given to researchers by the permission of the TKU Archive. Some indexes, registers and the documentation of the material in DDI model can be made available on the Internet for everyone to use and search.

The digitisation of the material of the Saami Folklore research project began autumn 2000 and is expected to take two years to complete. At the same time, the digitisation of all the material of the archives of ethnology, folkloristics and comparative religion has started.

In practice, the digitisation project started with the scanning of photographs, slides and transcriptions of the research material. The digitisation of the audio material started in the summer of 2001. The photographs, slides and transcriptions are scanned in two different versions. For the net the scanning resolution is 72 Dots Per Inch (DPI) and the pictures are saved in Joint Photographic Experts Group (JPEG) format. The pictures are sufficiently clear to be studied easily in a WWW environment, but they cannot be printed or published. The second version of the pictures is scanned to the archive in Tag Image File Format (TIFF) and their resolution is 300 DPI. The Archive version will also be copied on to two different CDs — one for use and as a copy of the material and one as a security copy. The Archive version of the picture, as well as the transcription, is sufficiently accurate for use in publishing, exhibitions and for other occasions2.

Problems to be solved.
In addition to the technical solutions there is much food for thought in the area of research and archive ethics. One of the main problems is the protection of the confidentiality of informants. Rules and solutions for this problem should be discussed together with other researchers and archivists who are concerned with these same problems. The archive databases belong to the area of legislation concerned with personal data (HetiL 523/1999; JulkL 621/99), because they include data about the informants, for instance, their worldview or religion that should remain confidential according to the law. So it is extremely important that the collected information not be used for purposes other than those for which it was collected: research and education. We also must remember that the archive material cannot be used for commercial purposes.

The fast development of information technology has created new ways of using and disseminating the archived material; the ways are so new and so different from the old ones that it has been impossible to prepare for them during the collecting of the material. The informants may have been afraid that the interviews would be broadcast, for example, but global dissemination of the material via the Internet is now a possibility for which it has been impossible to prepare.

Cooperation among traditional archives and researchers working with archived material is, and should be, extremely important. There is an urgent need for common rules and common procedures among those traditional archives that deal with digitised material. In principle, the names of the informants cannot be used for reasons concerning both legislation and research ethics. On the one hand, when we are dealing with older projects, such as Saami folklore, the original names of the informants have been used in all the articles and research. So it may be frustrating for the researcher for the Archive to start hiding facts that have previously been made available. On the

other hand, the informant also holds 'copyright' for his or her narratives. It is the product of the informant and his or her community; they might want to be recognised and remembered by it. In such cases, permission for use of the informant's name should be directly requested.

When the digitised archive material is given to the researcher, in other words, when the researcher gets the username and keyword of the database, or the CD which contains the copied archive material, he or she must formally agree to use the material only for commonly agreed purposes and also not to make the material available to another person or institution. When new interviews are made, documented permission for archiving and use of the material should be made, but when we are handling old material, responsibility for the protection of the informant lies with the researcher and the archive.

One of the main questions, we usually encounter is, of course, money. When starting the digitisation project, a primary issue was the provision of external funding for the project. But we have to keep the future in mind, too. When we use vast amounts of money for digitising our archive material, we may not want to surrender it and copy it for every possible researcher in the world at no cost. But is it ethically correct to collect money from colleagues? How much are they prepared to pay to consult the research material? How can you put a price on national cultural heritage?

Sources
a) Unprinted sources

The Project 2000. Kansallisen kulttuuriaineiston digitointi ja on line -käyttöön saattaminen. Hankesuunnitelma. [The digitising and on line -usage of the national cultural heritage. Project plan.] The Department of Cultural Studies at the University of Turku.

[http://www.utu.fi/hum/kultut/hanke/index.html]

The Pilot Project 2000. Saamelaisaineiston digitointi - saamelaisen folkloren tutkimus. Kulttuurien tutkimuksen laitoksen arkistojen digitointiprojektin pilottihanke. [Digitising the Saami folklore material - researching the Saami folklore. The pilot project of the digitising project of the Department of Cultural Studies.] The Department of Cultural Studies at the University of Turku.

[http://www.utu.fi/hum/uskontotiede/COLLCARD/ pilotti.htm]

TKU/A/00/79. Interview of Professor Lauri Honko 16.6.2000. Interviewers Pasi Enges and Tiina Mahlamäki. Audio collections of the TKU Archive.

b) Bibliography

Arkistolaki 831/94. [The Act on Archives]

Henkilötietolaki 523/99. [The Personal Data Act]

Herranen, Gun & Lassi Saressalo (eds.) 1978. A Guide to Nordic Tradition Archives. Turku: Nordic Institute of Folklore.

Honko, Lauri 1983: Folkloristin keräilytaloudesta. [Folklorist as a collector] - Kotiseutu 4/1983: 170-175.

Honko, Lauri 2000 (ed.): Thick Corpus, Organic Variation and Textuality in Oral Tradition. Studia Fennica Folkloristica 7. Helsinki: Suomalaisen Kirjallisuuden Seura, 3-28.

Huttunen, Hannu-Pekka 1992: Data-projektarkivet "ABBA" - en TRIP databas. - Herranen, Gun (red.): Data och folklore. NIF:s seminarium om databaserad registerhantering och kommunikation. NIF Rapporter nr. 7. Åbo: Nordiska instituitet för folkdiktning, 18-23.

Huttunen, Hannu-Pekka & Jaakko Isojunno & Marianna Kentala 1991: ABBA:n alkeet. TKU-arkiston ATK-ohjelma. Turku.

Kurkela, Vesa 1999: Äänitteiden säilyminen on keräilijöiden varassa. [The preservation of the Sound Recordings depends on private collectors] Helsingin Sanomat 15.2.1999.

Laki viranomaisen toiminnan julkisuudesta 621/99. [The Act on the Openness of Government Activities]

Mahlamäki, Tiina 2000. Äänitearkisto-opas. TKU-arkisto. [A guide to the Sound Archive] Uskontotieteen toimitteita 3. Turku.

Mahlamäki, Tiina & Pasi Enges 2001 (in print) From the field to the net: Analysing, cataloguing and digitising the material of the Saami folklore project. - Ulrika Wolf-Knuts (ed.) Input & Output. The Process of Fieldwork, Archiving and Research in Folklore. Turku: NorFa.

Nyberg, Patricia & Marjut Huuskonen & Pasi Enges 2000: Observations on Interview in a Depth Study on Saami Folklore. - Honko, Lauri (ed.): Thick Corpus, Organic Variation and Textuality in Oral Tradition. Studia Fennica Folkloristica 7. Helsinki: SKS, 489-536.

Rajamäki, Maria 1989: Introducing COLLCARD. NIF Newsletter 4, 35-39.

Saarinen, Jukka 2000: Det digitala arkivet och Internet. [Digital Archive and the Internet] - Ulrika Wolf-Knuts (ed.)

Vägen till Arkivet. Åbo: Åbo Akademi, 119-132.

Saressalo, Lassi & Huttunen, Hannu-Pekka 1985: Äänitearkisto-opas. [A Guide to the Sound Archive] Turku.

Suomen arktisen tutkimuksen nykytila ja strategian suuntaviivoja 1998. [The present state of the arctic research in Finland] Kauppa- ja teollisuusministeriön neuvottelukuntaraportteja 4/1998.

1 The village of Talvadas is situated quite near of the village of Outakoski.

2 You can read more about the technical details of digitisation from the Archives website.

* Paper presented at the IASSIST/IFDO conference 2001 in Amsterdam. Tiina Mahlamäki, Department of Cultural Studies, University of Turku, Finland.