Storage Management in the Field

This guide describes the aspects that have to be taken into account when handling large media files in the field where only a notebook is available.

General Framework:

- Modern field work includes the recording and immediate annotation of media files (audio/video) in the field. Media files, in particular video files, require much storage capacity.
- Current day light-weight notebooks that are typically used in the field are equipped with hard discs from 40 to 80 GB, i.e. the available storage capacity is limited. A strategy for storage management therefore is required.

1. General Remarks

The workflow has changed from central media digitization/capturing to remote operation, since this allows the researcher to start already in the field with studying the recordings and creating for example transcriptions. The general recommendations are to digitize audio with linear PCM and a sample frequency of 48 kHz and 16 bits. For video the working format in the field typically is either MPEG1 or MPEG4, since MPEG2 and especially DV streams take too much disk space. There is robust and reliable software available to digitize audio and to capture and convert DV streams with a notebook. To not take any risks use the recommended software and carry out early checks with your technical people. Smart solutions have to be found to handle the huge amounts of storage capacity required in the field, however.
For all you are doing don't forget that digitization and capturing take real time and that converting DV to MPEG1/4 also takes about real time. Both steps also will require power.

2. Audio/Video Storage Requirements

A stereo audio recording with 48 kHz/16 bit linear PCM of an hour will produce 700 MB of data – the capacity of a CDROM. A DV file of an hour which you first have to store on the notebook will produce about 16 GB. A conversion to MPEG1 for example would require additional 700 MB. Sometimes the generation of a separate audio stream to a wav file is required which requires additional 700 MB. After the conversion the DV file can be deleted, i.e. only for the conversion time you will need about 18 GB storage capacity for an hour recording. When you are using time markers and only capture shorter relevant sessions then only a fraction will be necessary. In all cases you have to write down the begin and end times carefully and communicate them with your archiving people.

3. Possible Solutions

Here a number of options are given. It depends on the field circumstances to determine the one that fits best:

- One smart method to reduce the requirements is therefore to process shorter sessions if possible.
- external large hard discs (>200 GB) – require power, bring additional weight and their capacity is also limited
- use CDROM burning – burning requires power, but is a good option, requires a number of empty CDROMs
- new notebooks may have DVD burners, but they require even more power; requires a number of empty DVD-R

Peter Wittenburg, MPI for Psycholinguistics, 12.4.2005.