DGS Corpus & Dicta-Sign: The Hamburg Studio Setup

Thomas Hanke, Lutz König, Sven Wagner, Silke Matthes

Institute of German Sign Language and Communication of the Deaf
University of Hamburg, Germany

THE AIM

Within the framework of the DGS Corpus as well as the Dicta-Sign project sign language corpora are being compiled. In a first step language data are being collected from Deaf informants who are invited to our studio and asked to conduct a variety of tasks presented to them. For this purpose, the DGS Corpus project and the Hamburg part of Dicta-Sign are working in close cooperation, sharing materials and experiences as well as using a joint studio.

Not taking into account budget restrictions, the setup of a sign language studio always is a balancing act between high quality recordings on the one hand (in order to ease the transcription process and possibly enable automatic processing), and on the other hand an environment where the informants still feel comfortable enough so that the recording situation does not have too much impact on the signing.

In the case of the DGS Corpus project, an additional constraint is that the studio is to be relocated several times as it was decided to make the recordings directly in the different regions of Germany. One of the implications of this approach is that the studio is operated by non-specialist deaf fieldworkers with limited time available for training.

THE MODERATOR'S ROLE

The fieldworker of a certain region also serves as the moderator during the elicitation sessions. In order to avoid influences on the language production it is crucial that the moderator is a Deaf person and that no other (hearing) person is present in the studio throughout the session. The moderator is responsible for the presentation and explanation of the individual tasks, the time management as well as a smooth run of the whole session. S/he observes the conversation, but only interferes with the conversation if absolutely necessary.

MODERATOR TRAINING

Leading the elicitation sessions and taking care of every aspect required leaves a heavy responsibility with the moderator. Training sessions are therefore needed to fulfil this task. However, with limited time available prior to the elicitation and especially with long elicitation sessions to be performed (seven hours for the DGS Corpus elicitation), a (semi)automatic control of the session should be implemented wherever possible. The custom software "Session Director" was therefore developed in order to support the moderator in his/her work and to ensure a smooth work flow.

While the Session Director user interface is quite straightforward to use, especially time management for the data collection sessions remains a demanding task and requires that the moderator is familiar with the program under all conditions. Moderators are introduced to the program within the fieldworkers' training courses.

## PRESENTATION OF THE TASKS

Each task is briefly introduced by the moderator, followed by a detailed explanation and instruction for the informants, presented in a DGS video clip on the informants' monitors. While these pre-recorded explanations ensure that all informants get exactly the same information and that nothing is left out, further clarifications given by the moderator might become necessary for some of the informants before starting a task.

The materials used as stimuli during a task comprise of different media formats, including pictures, drawings and video clips. They are shown as slides on the informants' monitors in a semiautomatic presentation, partly one slide following the other at a fixed speed, partly controlled by the moderator (e.g. allowing for in-between questions by the informants). Depending on the individual task the presentation of this material might be identical or different for the two informants.

THE STUDIO SETUP

The elicitation setting for both DGS Corpus and Dicta-Sign involves two informants interacting in different ways with each other and a moderator leading the session. During a recording session the pair of informants is sitting facing each other at an approximate distance of three meters. Camera positions in the studio guarantee that each informant is being filmed separately while there exists another camera to film the overall scene. The camera setup consists of seven cameras altogether:
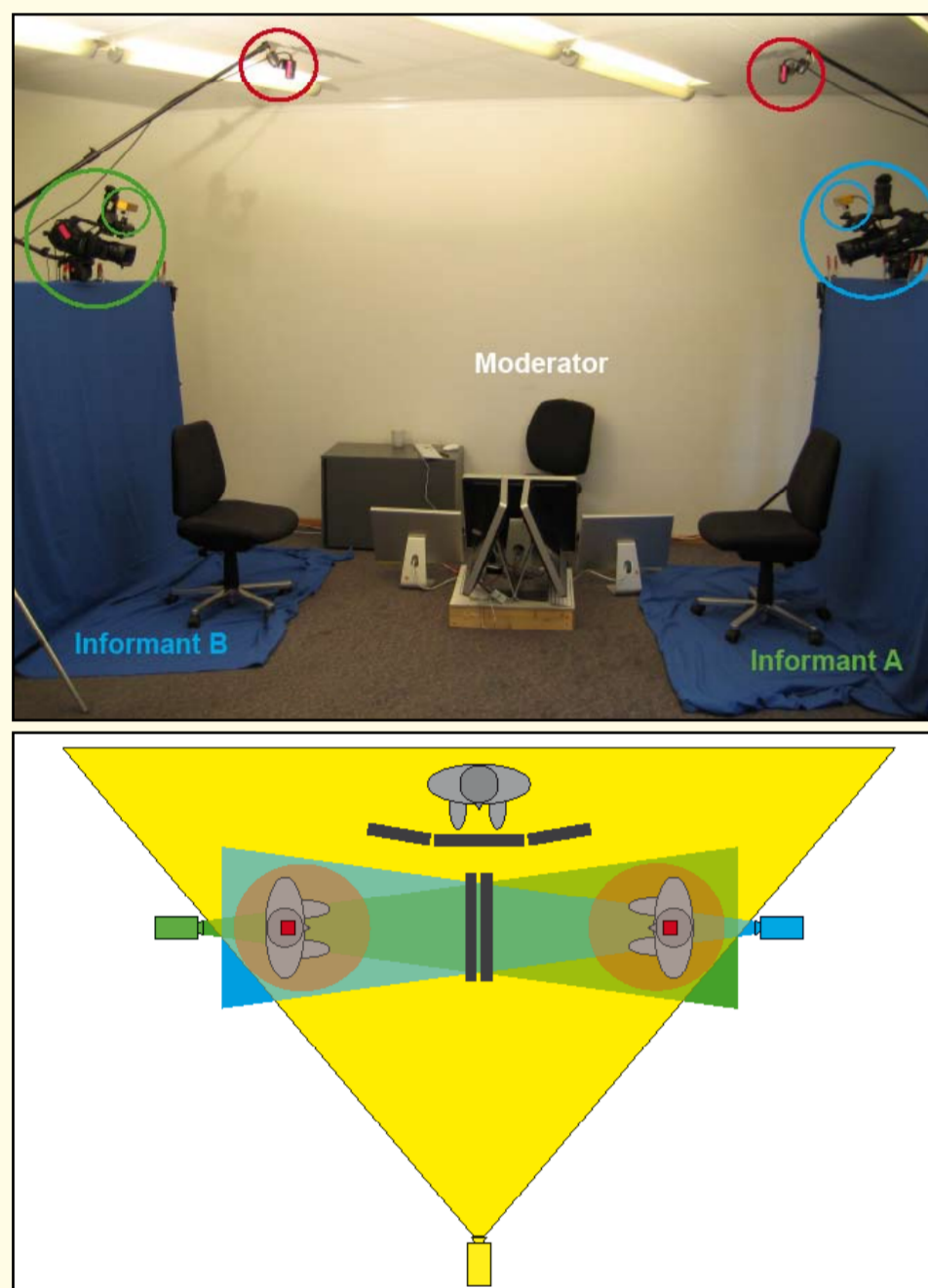
- Two HD cameras provide frontal views of each of the informants.
- Two stereo cameras, mounted on top of the frontal-view cameras, are to provide footage that allow image analysis to reconstruct 3D information and help automatic processing.
- Two birds-eye cameras capture each informant's signing from above to help human transcribers to interpret the signing.
- One HD camera captures the whole scene to give the transcriber a quick overview and to help him/her to exactly identify interactions between the three participants.

Elicitation material to be shown to the informants is presented on screens in the middle between the signers very close to the ground in order not to interfere with their views of each other. The moderator of a session is sitting next to the informants. Monitors in front of the moderator display the "Session Director" as well as what can currently be seen on the informants' screens.
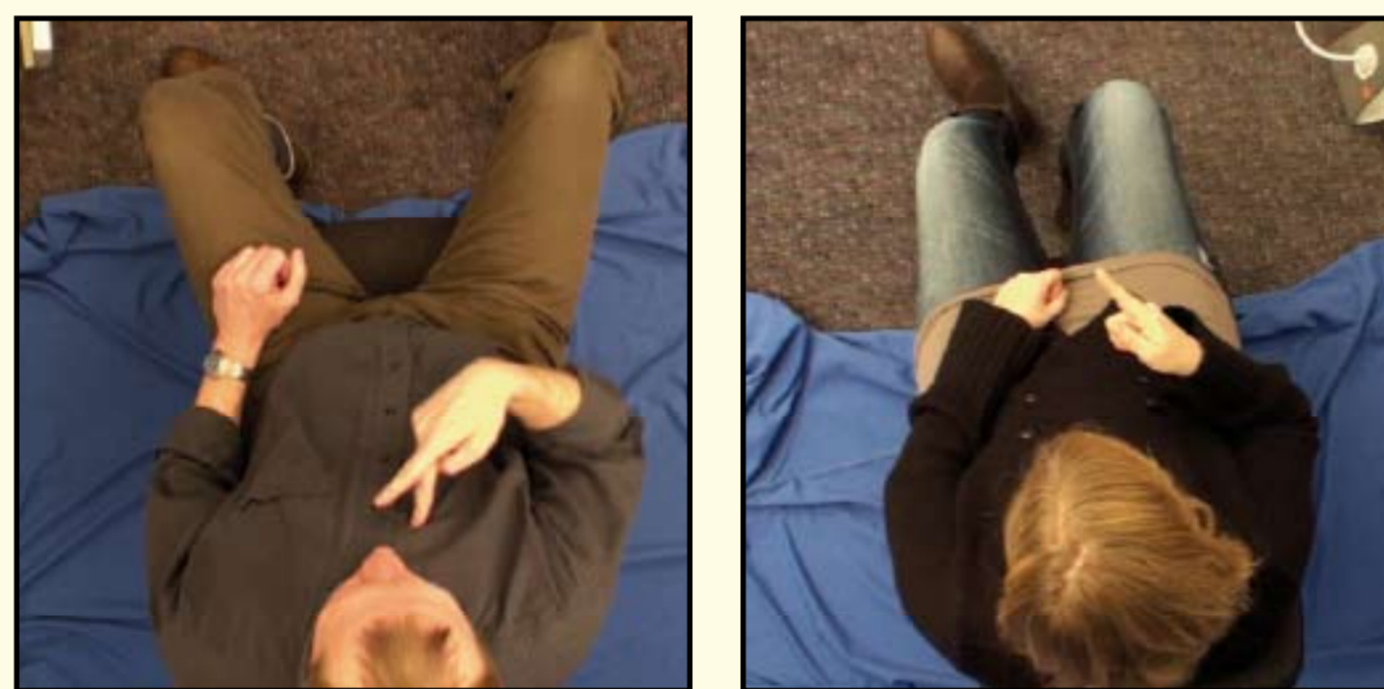

POSSIBLE CAMERA POSITIONS

A variety of different seating arrangements and camera positions were taken into account and tested for the Hamburg studio:

In earlier projects the informants were seated opposite each other but at a slight angle. The cameras were positioned at eye height on the side of each informant. The drawback of this approach is that some informants constantly target their signing back and forth between the addressee and the camera, which makes it difficult to identify important aspects as body shifts in the discourse situation.
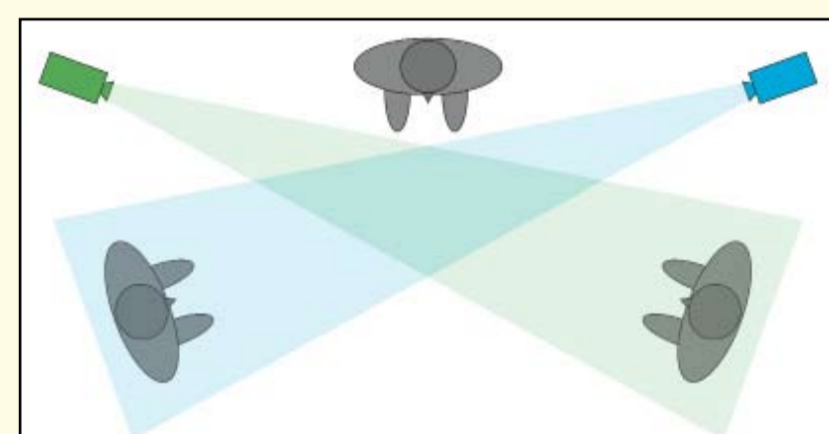

Instead, a setting where the informants are directly facing each other was desired in order to reduce unwanted body shifts as well as to assure a relaxed conversational situation. Leaving the cameras in a side position however provides a view on the informant at eye height but at a side angle, which again makes body shift and eye gaze tracking difficult for the transcriber as well as for semi-automatic processing.
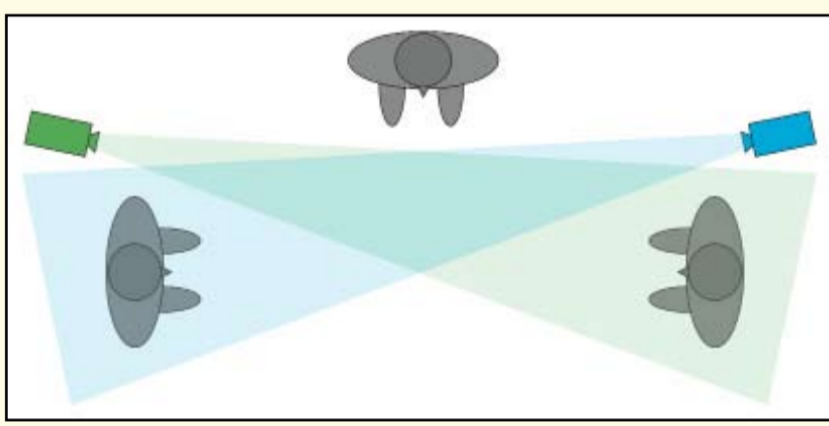
A setting with the main cameras positioned above and behind the heads of the informants, as decided for the Hamburg studio, provides a view from the front but with a slight angle from above. However, pre-tests revealed that with a distance of approximately three meters, the distortion introduced by the elevated position of the camera does not negatively affect the transcription from video. Instead, this setting provides a front view of the informant similar to the addressee's, allowing identification of body shifts as well as eye gaze direction more easily. At the same time, especially with the monitors being located on the floor between them, the informants did not feel this to be an unnatural distance for their interaction.

SESSION DIRECTOR

A custom software, "Session Director" allows the moderator to present slides to the informants by the click of a button, and to keep track of the time elapsed for each individual task as well as the whole session. While one screen facing the moderator displays Session Director, two additional screens exist to observe the presentations shown on the informants' screens.
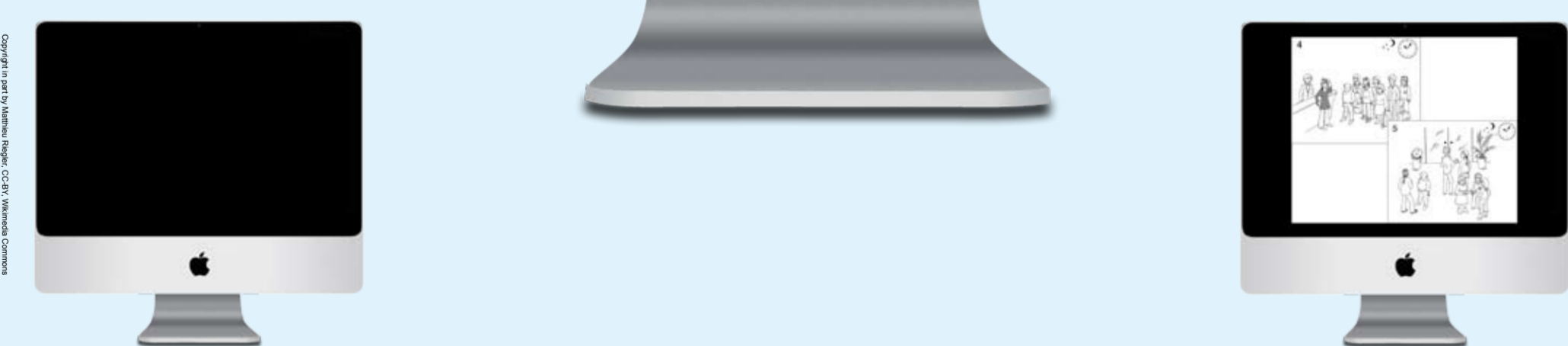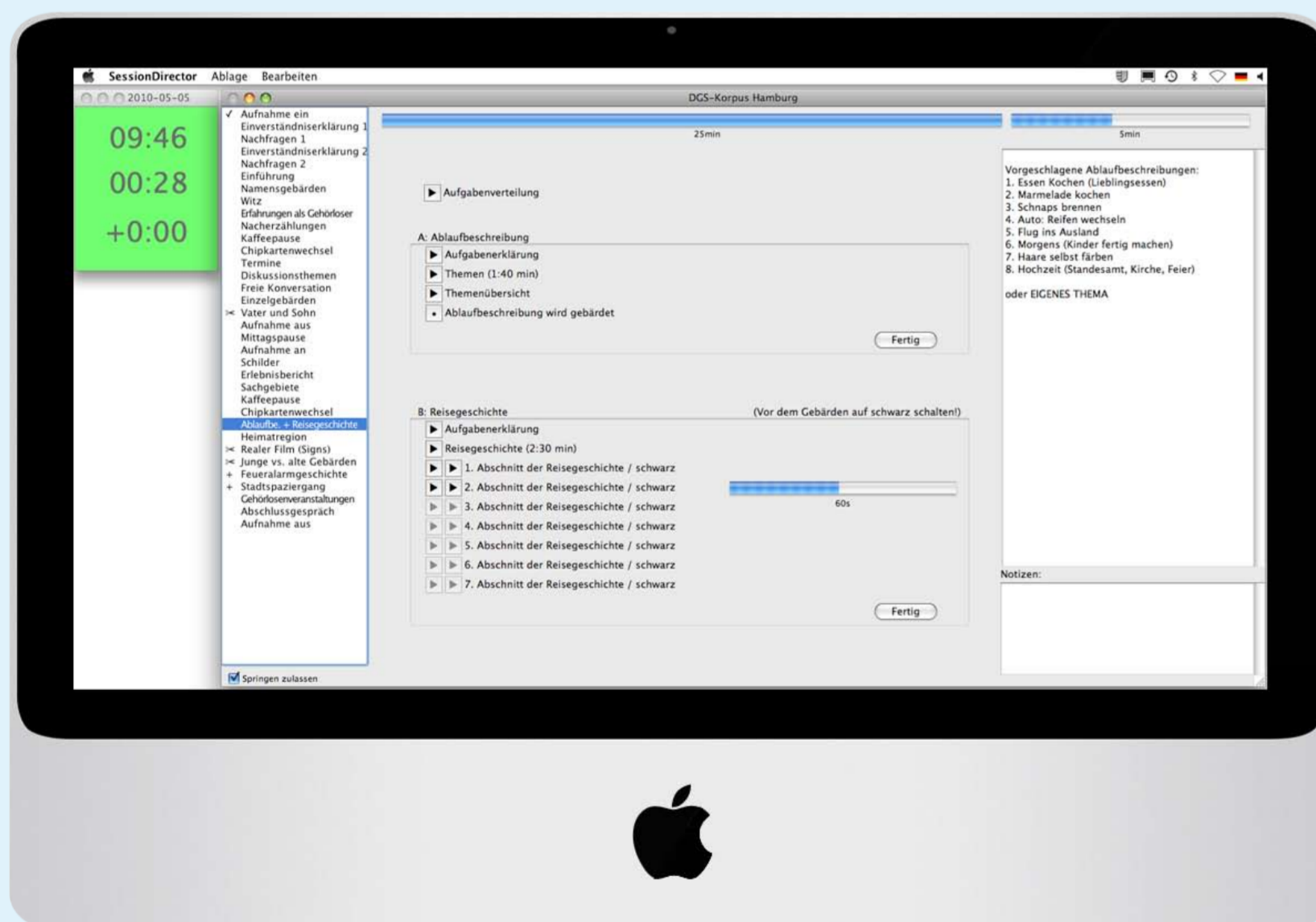
Session Director's main window shows a list of all tasks as well as details of the task currently worked on. This includes a progress bar showing the time already spent on the task in relation to the time planned as well as the sequence of subtasks (such as the task explanations, stimuli presentations and conversations between the informants). While the sequence of events is predefined, it is the moderator clicking start buttons on the screen to activate the next step. This allows the moderator to check if all explanations have been understood or further clarification is needed before moving on.

While it is neither doable nor desirable to keep each individual task to the planned duration, the total session time should be kept close to the plan. For this purpose, Session Director shows another window giving the actual time, the elapsed time in the session, and the time before/behind plan. The window changes colours to signal when deviations exceed a certain threshold and require corrective action.

SESSION DESCRIPTION FILES

When launched, Session Director loads an XML file describing the session. For each task and subtask, it defines the expected and maximum duration, the text of the user interface elements visible in Session Director and of course the ids of the slides from an Apple Keynote document to be shown to the informants, either as a common set or separately for each informant. Furthermore, it defines the relative importance of each task which Session Director will eventually use to mark tasks that can be skipped. In addition, text can be entered that will be displayed alongside with the task detail.

In general, the order of tasks, including breaks, is pre-determined by the XML session description. The moderator has, however, the freedom to rearrange tasks, to change the expected duration of the session Session Director measures progress against or to halt the task time with a pause function should a spontaneous break become necessary.
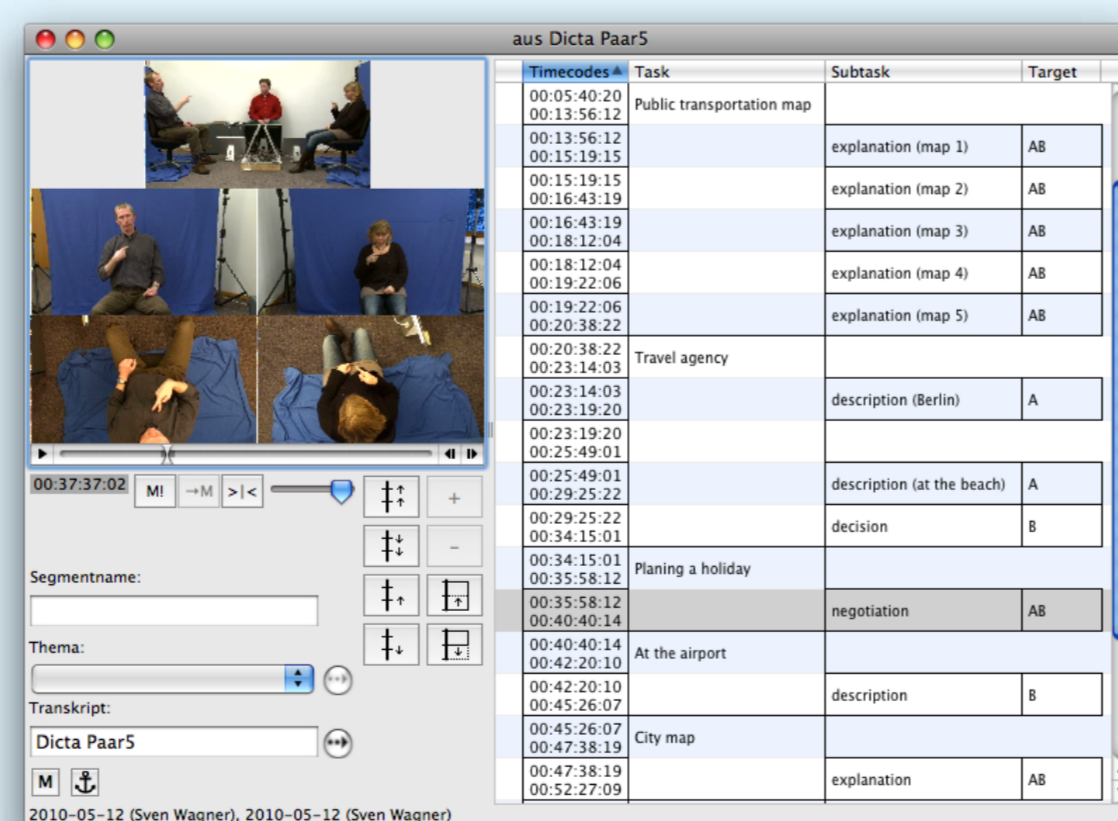
All the moderator's interactions with Session Director are logged with time stamps. This allows us to determine automatically when on the videos certain tasks (or pauses) can be found and also to conclude from our knowledge of the tasks who presumably is the active signer at a given point in time. Notes that the moderator takes during the session are also output into that log file.

Session Director is available free of charge for MacOS X only as it heavily relies on MacOS-specific functionality. Source code is available upon request.

MOVING THE STUDIO

While for Dicta-Sign data collection is being conducted for four different sign languages, organised seperately by the project partners in their respective countries, the DGS Corpus project aims at collecting data from different regions in Germany. In order to avoid dialectal mixing it was decided to make the recordings directly in the regions instead of inviting participants to one central place. The studio therefore needs to be relocated twelve times over the course of two years.

A room is needed of at least 5m x 5m, optimally with a ceiling height of at least 3m, as well as a smaller room next door for the technician. Variable equipment exists to adapt different room layouts (e.g. different camera racks).

For the relocation of the studio, transport cases exist for all equipment including spare parts. Deinstallation, actual transportation and installation at the next site is organised by an external service provider. Approximately one day needs to be scheduled for assembling and disassembling respectively. After the installation of the mobile studio our Deaf technician checks the functionality prior to the first elicitation session.

TECHNICAL DETAILS

The HD cameras we use are 3 Sony EX3 and 2 Sony HRX-MC1P (for the birds-eye views), all recording at 1080i25. These cameras store data locally on memory cards (only used as a fallback solution) and at the same time stream into MacPros running FinalCutPro.

The stereo cameras are PointGrey Bumblebee 2 models capturing 640x480p48 for each channel. They are connected to MacPros running capture software provided by Dicta-Sign partner University of Surrey (currently under Windows XP).

With total data rates of 700 GBytes per hour, we are not able to transfer the data to the Hamburg server before the next session starts. Instead, we swap hard disks (Raids consisting of two 2TBytes hard disks for each computer) after each fourth session and transport them back and forth in special suitcases. A local backup is available just in case.