ELDP Language documentation training

February 16-19, 2015

The ELDP training workshop will provide a variety of formal, informal and practical sessions dealing with topics in language documentation focussing on video recording and language use, video conversion, annotation with ELAN, digital archiving, data management, metadata collection. This workshop will take a group-based "workflow" approach, proceeding from issues and practicalities in collecting primary materials (video), to analysing and archiving them.

The use of video in linguistic research and language documentation

Language use is fundamentally multimodal. When we speak point to places, we nod to signal our understanding and we outline our thoughts with our hands. Moreover we coordinate our actions with our conversational partners. The way we talk, the words we use, the information we include depends on the common ground we share with our interlocutors and different social and contextual dimensions. This complexity of language use is best captured with video. Language documentation projects as well as any linguistic research project investigating dimensions of conversational language use should record video. However, video training is often limited to the technicalities of how to use a video camera and microphones but does not provide a theoretical framework for the features of language use. What is missing is an understanding of the social, linguistic and cognitive foundation of language use so that it becomes clear what needs to be captured how.

In this workshop, we will address these topics and participants will learn about the theoretical foundation of language use as a joint activity. Focussing on manual gesture participants will learn how gesture data play an important role in understanding the fundamentally multimodal nature of language use. We will provide and discuss detailed guidelines of how to video record and this will be trained with hands-on practice. Participants will collect natural language use data of different types and learn how to use video equipment. This will be followed by a review of the recordings and guidelines for improvements will be discussed. We then will train participants in how to convert the video recordings for working on computers and annotation in ELAN.

Transcription and annotation with ELAN

ELAN is a transcription and annotation tool developed at the Max Planck Institute for Psycholinguistics in collaboration with linguists over the past 15 years. The tool is free and widely used by linguists (https://tla.mpi.nl/tools/tla-tools/elan/). ELAN allows for easy and quick transcription and translation, glossing, annotation and coding of linguistic and behavioural categories. The data can be searched across files. In the training participants will learn how to set up and use ELAN as a transcription and annotation tool. ELAN files will be set up and participants will learn how to transcribe in ELAN, how to search across files and how to export the data into a data base for mining.

Archiving

Today digital archives preserve and publish collections of audio and video recordings and texts. The materials linguists are collecting are in times of dramatic decline of linguistic diversity invaluable documentations of our cultural and linguistic heritage. Therefore the corpora should be digitally archived and made accessible to other researchers and the speaker community themselves. Fieldworkers collect large amount of materials, which needs careful management and detailed metadata to create a corpus that can easily be archived and discovered by user worldwide. This training will teach the basic principles of data management for archiving. We will also introduce metadata editors which make record keep systematic and straight forward.

PROGRAMME

	Monday	Tuesday	Wednesday	Thursday
9:30-11:00	Visual mode of Language	Video Review/ Transfer & Conversion	Metadata	Clinic
11:00-11:30	break	break	break	break
11:30-1:00	Visual mode of Language	Video Review/ Transfer & Conversion	ARBIL	Clinic
1:00-2:00	lunch	lunch	lunch	lunch
2:00-3:30	How to record video	ELAN	Workflow and software integration	Clinic
3:30-4:00	break	break	break	
4:00-5:30	Equipment selection & recording practice	ELAN	Data management	