

Hidden languages in a digital world: the case of sign language archives

Pietro Braione (Università di Milano-Bicocca, CINI), Carlo Cecchetto (CNRS, Università di Milano-Bicocca), Carlo Geraci (IJN, CNRS, DEC, ENS), Roland Pfau (University of Amsterdam), Josep Quer (ICREA, Universitat Pompeu Fabra)

Background. Video technology has been widely available for several decades, and since the late 1990s also in digital format (Castells 2001). This technological breakthrough opened huge opportunities for the study of sign languages (SLs) in all domains of social sciences (Kusters et al. 2017), which in turn speeded up the process of recognition of local SLs by national governments (DeMeulder et al. in press). The video materials document the oral memories of visual languages (i.e., SL).

The project. The (EU funded) Sign-Hub project aims at providing a digital platform to archive and give access to: on-line SL grammars, SL assessment tools, SL atlas of linguistic structures, and life narratives of elderly signers documenting the history of Deaf communities. Sign-Hub is generating a huge amount of video data that are and will be used in various ways.

Goals. We focus here on the following topics:

- Opportunities and challenges of use and re-use of historical narratives for language documentation and culture preservation purposes
- Risks and challenges associated with long-term preservation and streaming of video materials and their metadata

Discussion. The Sign-Hub platform is designed to fulfill two requirements: it must 1) provide content creators (grammar writers, atlas compilers and SL historians) with a user-friendly interface to upload and manage their contents (Fig. 1); and 2) offer a pleasant experience to the final users (a.o., linguists, students, historians, Deaf people, general audience, etc.).

As a case study, we discuss **life narratives of elderly signers**. In the course of the second half of the 20th century, several events impacted the lives of deaf people and the destiny of European Deaf communities. Some are part of the World history (e.g., WW2, Shoah, Civil Wars or Cold War). Others are specific to Deaf communities (e.g., experiences in Deaf institutes, situation of ethnic, or gay/lesbian minorities) or the lives of Deaf individuals (e.g., first encounter with a deaf person, cases of language discrimination).

Data for this part of the project come from already existing recordings (for France and Israel) or from newly recorded interviews (for Germany, Italy, Netherlands, Spain, Turkey). As for the former, we provide: 1) a digitization of old videotapes, 2) subtitling and voicing in the local language, 3) English subtitles. As the latter, we provide 1) subtitling in the local spoken language and in English, 2) partial linguistic annotation in ELAN (Crasborn & Sloetjes 2008; see Fig. 2), 3) Metadata in CMDI format specific to SL (Crasborn 2017).

While the memory of 20th-century events from the perspective of Deaf communities will be preserved, the very same material will also be used as:

- documents about the history of Deaf communities
- samples of the language of elderly signers
- the basis to study linguistic change across Deaf generations
- material for other social sciences (e.g., anthropology, sociology)
- materials for documentaries for the general public

Digital format is extremely perishable and subject to fast technological changes. We are faced with a double challenge: 1) archiving the video materials for long-term preservation; 2) making them accessible via web-interfaces. Our solution is to use a double storage policy: the same video is generated in two formats: a high quality format is stored in public archives for long-term preservation; a compressed form is used for streaming and immediate access to content.

Figure 1: Screenshot of the workspace for content creator

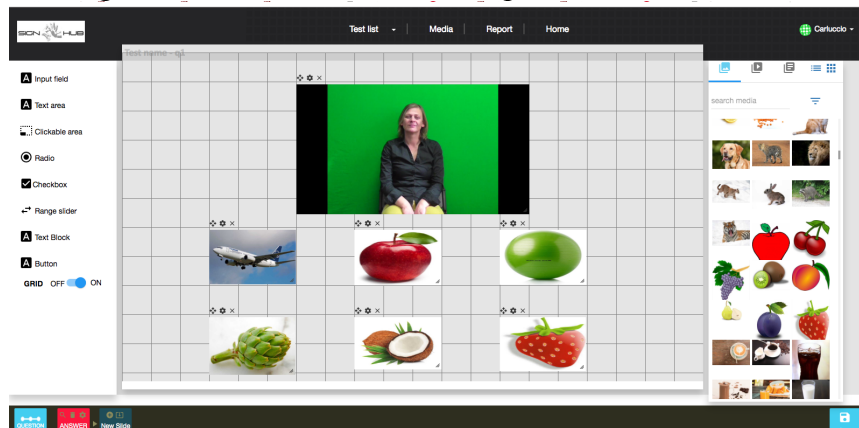
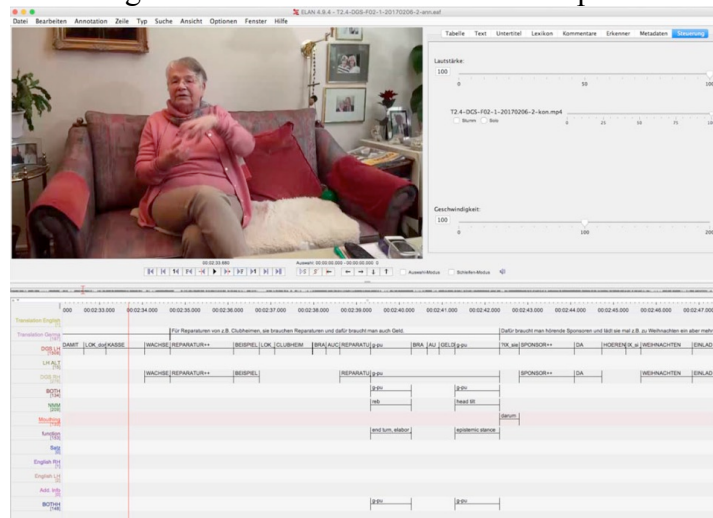


Figure 2: Screenshot of ELAN workspace



References

- Castells. 2001. *The internet galaxy*. OUP.
- Crasborn, & Sloetjes. 2008. Enhanced ELAN functionality for sign language corpora. In: LREC 2008.
- Crasborn. 2017. SL-TL: Sign language data archived at The Lang. Archive.
- Kusters et al. 2017. *Innovations in Deaf Studies: The Role of Deaf Scholars*. OUP.
- DeMeulder et al. In press. *Recognizing Sign Languages*. Multilingual Matters.