Speech, Voice, Text, and Meaning

A Multidisciplinary Approach to Interview Data through the use of digital tools

Arjan van Hessen
University of Twente
The Netherlands
a.j.vanhessen@utwente.nl

Silvia Calamai
Università di Siena
Italy
silvia.calamai@unisi.it

Henk van den Heuvel
Radboud Universiteit Nijmegen
The Netherlands
h.vandenheuvel@let.ru.nl

Stefania Scagliola
University of Luxemburg
Luxemburg
stefania.scagliola@uni.lu

Norah Karrouche
Erasmus Universiteit Rotterdam
The Netherlands
karrouche@eshcc.eur.nl

Jeannine Beeken
University of Essex
United Kingdom
jeannine.beeken@essex.ac.uk

Louise Corti
UK Data Archive
United Kingdom
corti@essex.ac.uk

Christoph Draxler
Ludwig-Maximilians-Universität München
Germany
draxler@phonetik.uni-muenchen.de

ABSTRACT

Interview data is multimodal data: it consists of speech sound, facial expression and gestures, captured in a particular situation, and containing textual information and emotion.

This workshop shows how a multidisciplinary approach may exploit the full potential of interview data. The workshop first gives a systematic overview of the research fields working with interview data. It then presents the speech technology currently available to support transcribing and annotating interview data, such as automatic speech recognition, speaker diarization, and emotion detection. Finally, scholars who work with interview data and tools may present their work and discover how to make use of existing technology.

KEYWORDS

interview data; speech processing; emotion detection; transcription; annotation; NLP

1 Introduction

As increasingly sophisticated new technologies for working with numbers, text, sound and images come on stream, there is one type of data that begs to be explored by the wide array of available digital humanities tools, and that is interview data [2, 4] (Fig. 1). A series of four workshops from 2016 to 2018, held in Oxford, Utrecht, Arezzo and Munich, brought together scholars from different research fields to explore their needs, requirements, and the current state of the art with respect to interview data.
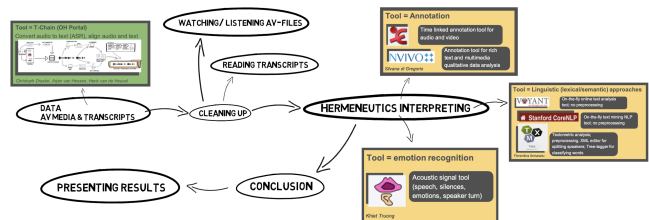

Figure 1: High-level simplified journey of working with interview data

When considering research processes that involve interview data, we observe a variety of scholarly approaches, that are typically not shared across disciplines. Scholars hold on to engrained research practices drawn from specific research paradigms and they seldom venture outside their comfort zone. The inability to 'reach across' methods and tools arises from tight disciplinary boundaries, where terminology and literature may not overlap, or from different priorities placed upon digital skills in research. We believe that offering accessible and customized information on

how to appreciate and use technology can help to bridge these gaps.

2 The Workshop

This workshop aims to take stock of our efforts to break down disciplinary boundaries by offering scholars the opportunity to apply, experiment and exchange digital tools that have been developed in the realm of Digital Humanities. We would also like to share our insights with a wider audience of scholars coming from different disciplines. Point of departure is a webportal that we have developed for automatic speech to text conversion (transcription) by Automatic Speech Recognition (ASR).

This half-day workshop is open for everyone interested in the processing of spoken narratives for research purposes (scholars) or for archiving and opening up interview data (archivists). It will provide cross-disciplinary knowledge exchange between all participants by presenting the following elements:

- Outline of the various approaches of different disciplines that work with spoken data and how they use annotation, text analysis and emotion extraction tools to process their data [6]
- Outline and hands-on components of the open source tool the T-Chain (Fig. 2). Participants can process their own audio-clips in Dutch, German, Italian or English. and explore how to further process the ASR output [7, 8]
- Connect the ASR-results to existing written text sources (books, newspapers, blogs)
- Short lectures by invited speakers
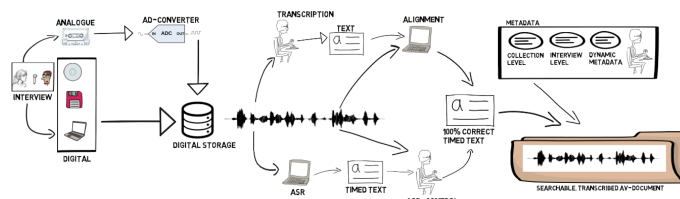- Figure out how to strengthen and expand the community around Interview Data and Technology

Figure 2: schematic view of the T-Chain: processing of spoken documents from recording/digitization into transcribed digital files including metadata

3 Workshop website and user participation

The workshop has a maximum of 25 participants. The publicity, interaction (questions, remarks, registration with special wishes/ questions concerning the workshop) with the potential participants, will go through the website: https://oralhistory.eu/workshops/icmi. Participants who are interested in contributing to the workshop, can submit their proposals via the same website.

The organizers will ask at least two workshop participants to give a presentation about their own working methods, experiences,

and wishes regarding the elaboration of the interview process. The selection will be made by the organizers.

4 Programme

9:00 – 9:30	Introduction and short presentation on 'Digital Humanities approaches to interview data - can historians, linguists and social scientists share tools?'
9:30 – 11:00	Preparing your audio-data, uploading the audio to the portal and automatic recognizing the speech. Correcting the ASR-results
	Downloading the (corrected) results and improving the readability
11:00 – 11:15	Coffee break
11:15 – 11:45	Invited speaker 1 incl. QA
11:45 – 12:15	Invited speaker 2 incl. QA
12:15 – 12:45	General discussion and close of meeting

REFERENCES

[1] Oral History and Technology Collaboration. 2018. Oral History and Technology group. Website: https://oralhistory.eu.

[2] L. Corti, N. Fielding, 2016. Opportunities From the Digital Revolution: Implications for Researching, Publishing, and Consuming Qualitative Research. SAGE Open. https://doi.org/10.1177/2158244016678912 .

[3] B. A. Lanman and L. M. Wendling, 2006. Preparing the Next Generation of Oral Historians: An Anthology of Oral History Education. Altamira Press.

[4] F. de Jong and A. van Hessen and T. Petrovic and S. Scagliola, 2014. Croatian Memories: speech, meaning and emotions in a collection of interviews on experiences of war and trauma, Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14).

[5] S. Scagliola and L. Corti, 2018. Oral History under scrutiny in München - Cross disciplinary overtures between linguists, historians and social scientists, CLARIN website Blog.

[6] K. P. Truong and G. J. Westerhof, and S. Lamers and F. de Jong and A. Sools, 2013. Emotional expression in oral history narratives: comparing results of automated verbal and nonverbal analyses, Proceedings of the Workshop on Computational Models of Narrative, CMN 2013, 2013 Workshop on Computational Models of Narrative. CMN 2013 Hamburg, Germany.

[7] H. Van den Heuvel and A. van Hessen and S. Scagliola and C. Draxler, 2017. Transcribing Oral History Audio Recordings – the Transcription Chain Workflow. Poster at the CLARIN EU Conference, Budapest, September 18/19- 2017.

[8] C. Draxler and H. Van den Heuvel and A. van Hessen and S. Calamai and L. Corti and S. Scagliola, 2020. A CLARIN Transcription Portal for Interview Data. Proceedings of LREC 2020, Marseille