

A Multidisciplinary Approach To The Use Of Technology In Research: The Case Of Interview Data

Arjan van Hessen (a.j.vanhessen@uu.nl), Utrecht University, The Netherlands and Stefania Scagliola (stefania.scagliola@uni.lu), C2DH, University of Luxemburg and Louise

1. Summary

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**. (De Jong, 2014), (Corti and Fielding, 2016). A stream of concerted activities by the authors of this workshop based on their engagement with interview data from different perspectives, led to a series of 4 workshops from 2016 to 2018, held in Oxford, Utrecht, Arezzo and München, that were funded by [CLARIN](#) (Common Language Resources and Technology Infrastructure).

This workshop intends to present the motivation for this DH2019-workshop by sketching the complex design and results of these series of multidisciplinary workshops. The premise was that the multimodal character (text, sound and facial expression) and multidisciplinary potential of interview data (history, oral and written language, audio-visual communication) is rarely fully exploited, as most scholars focus on the analysis of the textual representation of the interviews. This might change by getting acquainted with scholarly approaches and conventions from other disciplines.

2. Aim of this workshop

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.

This workshop aims to break down some of these barriers by offering scholars who work with interview data the opportunity to apply, experiment and exchange tools and methods that have been developed in the realm of Digital Humanities.

3. Previous work

As a multidisciplinary group of European scholars, tools and data professionals, spanning the fields of speech technology, social sciences, human computer interaction, oral history and linguistics, we are interested in strengthening the position of interview data in Digital Humanities. Since 2016 we have organized a series of workshops, supported by CLARIN on this topic (<https://oralhistory.eu/workshops>).

Our first concrete output was the development of the T-Chain, a tool that supports transcription and alignment of audio and text in multiple languages. Second, we developed a format for experimenting with a variety of annotation, text analysis and emotion recognition tools as they apply to interview data.

4. The workshop

The half-day workshop will provide a fruitful cross-disciplinary knowledge exchange session. It will:

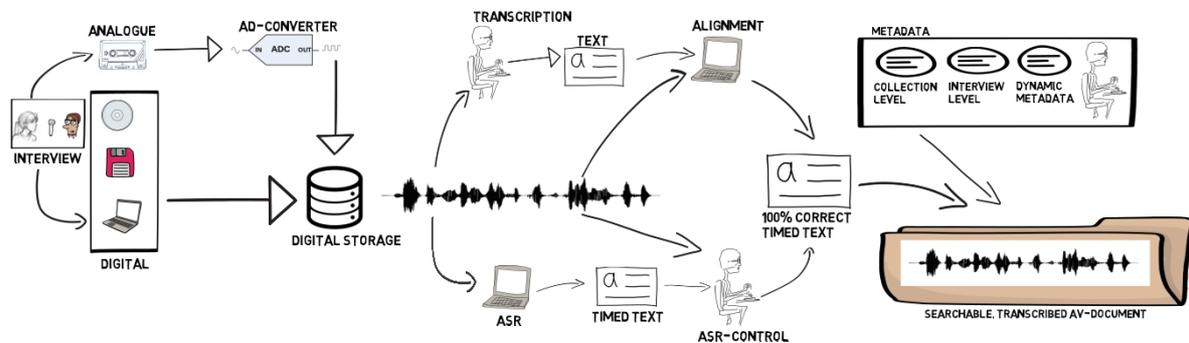
- Use presentations and hands-on sessions to explore annotation, text analysis and emotion extraction tools with interview data;
- Test the T-Chain with participants' own audio-clips;
- Attract open source Speech-to-Text software developers to expand the number of languages that could be integrated into the T-Chain.

4.1. Scope and organisation of the series of workshops

The organising team created a community of experts from The Netherlands, Great Britain, Italy and Germany who actively scoped and assembled invitations to scholars to the workshops. The countries were selected on the basis of the availability of mature open source speech recognition software, as the first goal was to develop a portal for automatic transcription and alignment of interview data for different

languages. Scholars and archivists who participated were those who work/teach with interview data, and who were interested in the use of technology to facilitate transcription and annotation of interview data and explore cross disciplinary analysis and interpretation of data. They represented the following communities:

- Historians and social science users who undertake research with recorded interview data sources;
- Linguists who use spoken language sources;
- Software tools specialists who develop support data processing and analysis tools.



The first three workshops thus focused on exploring user requirements and testing the performance of various speech-to-text software on interview data that was provided by researchers and data curators. This led to the development of the Transcription Chain (henceforth, T-Chain) a portal for automatic transcription and alignment of interview data in English, Dutch, German and Italian. This tool is meant to support the first phase of the research process, the transcription of interviews, and to create an open source easy to access web resource with different choices for the format of the output, anticipating the onward import into a variety of tools.

4.2. Cross disciplinary overtures in München

During the fourth workshop in München the scope was broadened to the subsequent phases of the research process: the annotation and analysis of the data. The presumption was that the multimodal character (text, sound and facial expression) and multidisciplinary potential of interview data (history, oral and written language, audio-visual communication) could be better exploited by bringing diverse approaches together and encouraging the uptake of digital tools. Anticipating that this diversity of participants and tools would make the organisation of the workshop complex, a careful design of the workshop was key to ensuring 'satisfying experiences' and countering 'disorientation'. To this end the following principles were applied:

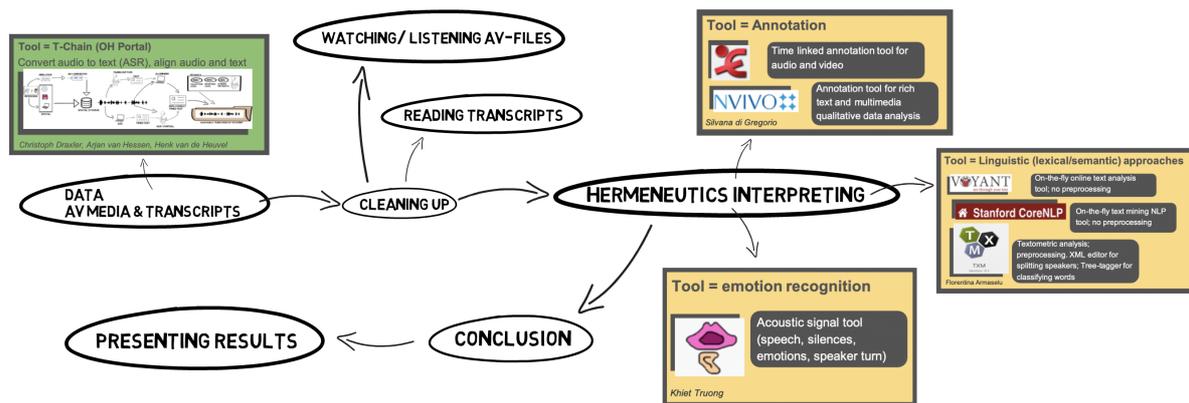
1. gathering detailed information on the participants prior to the workshop to tailor the sessions to their level of digital savviness,
2. collecting and preparing data that was familiar to the participants in both a common language (English) and in their native language,
3. building on homework assignments to install and become familiar with a number of tools,
4. making sure that during the workshop a participant with advanced digital skills was represented in each of the language groups, and
5. eliciting and recording feedback on use of the tools directly after the session exercises through group interviews.

The first session of the 3-day workshop was devoted to testing the first version of the T-Chain with German, Dutch, English and Italian data. In the subsequent three hands-on sessions participants worked with proprietary and open source annotation tools, common among social scientists, with text mining tools used by computational linguists, and with emotion recognition tools used by computer scientists. Prior to starting the hands-on sessions, it was deemed necessary to introduce, using jargon-free language, the range of diverse research profiles of the disciplines represented.

4.3. A parade of research trajectories

Each of these disciplines uses a different approach when working with recorded interview data, and within every discipline there exist distinct sub-disciplines. Rarely is there a unanimous voice on methods, analysis and use of tools. Even speaking of 'linguistics' is an over-simplification, just as the term 'oral history' is a broad term for a variety of approaches to interpreting interviews on people's life stories. For example, whereas an oral historian will typically approach a recorded interview as an intersubjective account of a past experience, another historian might consider the same source as a factual testimony. From a different perspective, a social scientist is likely to try to discover common themes and similarities and differences across a whole set of interviews. Thus, disciplinary approaches represent distinct analytical frameworks that might make use of tools in different ways. To illustrate the variety of landscapes, we invited workshop participants to consider 1-2 'research trajectories' that reflected their own approach(es) to working with interview data. This enabled us to come up with a high-level simplified journey and to identify how and where the digital tools might fit into the researchers' own workflow. The

diversity of practices of course implies that there are many variations of this trajectory.



High-level simplified journey of working with interview data

4.4. Tools for Transcription, annotation, linguistic analysis and emotion recognition

The researchers were invited to work in four 'language groups' of 5 to 6 people (Dutch, English, Italian and German) in hands-on sessions, using step-by-step worksheets and pre-prepared interview extracts. A distinction can be made between types of tools that support the research process, in the sense that technology substitutes manual labour, and those that have actual impact on the interpretation of the data. The T-Chain, developed with CLARIN support, with its speech to text and alignment software, can partly substitute the cumbersome transcription of interviews, a practice that is common to anyone working with interviews. The need for transcription is commonly understood across disciplines. Then there are tools that aid the annotation of text and audio-visual data by offering a structured system for attributing meaning to passages. At this point the common needs tend to decrease, as the choice of a particular annotation tool leads to an engrained practice of research, that requires time to build up and cannot be easily traded for an alternative. For this purpose, the open source tool ELAN was used, and compared with the proprietary qualitative data software NVivo. In the following two sessions the experimental component increased, as the same interview data was explored with computational linguistic tools of increasing complexity. The two tools used were Voyant and NLPCore, web-based tools, that allowed the processing of transcripts 'on the fly', enabling a whole set of specific language features to be directly identified. The second tool, T/M, had to be installed and allowed for a more granular analysis of language features, requiring the integration of a specific language model, the splitting of speakers, the conversion of data into computer readable XML language, and the 'lemmatization' of the data. The last session was the most daunting one, illustrating what the same data yielded when processed with the acoustic tool PRAAT, and the facial recognition tool Open Smile.

4.5. User experience and evaluation

A first analysis of the user experience seems to suggest that scholars are very much open to cross-fertilization, that the likelihood of a digital tool being used by a scholar significantly increases when the tool is transparent, and finally, that legal and ethical concerns are paramount in deciding whether or not to use a digital tool. First, there is great potential to cross-fertilize approaches. Oral historians and social scientists generally enjoyed trying out each other's methods, in particular analytic tools that might help complement their own 'content-driven' approaches to working with interview data, for instance to elucidate features of spoken language.

Second, knowing what a digital actually does 'behind the scenes' increased participants' sense of the tool's usability. Following that, tools must be flexible. Scholars are only willing to integrate a digital tool into their existing research practice and methodological mindset, if the tool can easily be used or even adapted to fit needs. Limited functionality of the free easy-to-use tools, and the observed methodological and technological complexity and jargon-laden nature of the dedicated downloadable tools, despite the availability of clear documentation, were both seen as significant barriers to use in everyday research practice. Specifically complex in this regard was the experience with the linguistic tools, especially those that required the data to be pre-processed. These have a high learning curve. While output triggers fascination and curiosity, it appears difficult to translate insights into the structure of language in an entire corpus, or in meaningful comparisons of subsets, to one's non-computational practice of interpreting individual personal accounts. The same applies to the emotion recognition tools. The messy data that pre-processing interviews about migration yielded, led to sifting out possible hypothesis, but not to a deeper understanding of the experience of migration. The real challenge lies in being able translate insights with regard to scale, frequency and paralinguistic features into the classic interpretation of the interview data. Often this means looking at other things, for instance the amount and character of silences within an entire corpus. This may reveal a pattern that is typical for this group of respondents or for this type of spoken data in general but is not related to migration. The same question about whether a shift in focus is needed to discover new practices applies to using speech to text software and alignment

of audio and text. The most salient conclusion of the exploration could be that the traditional practice of interpretation of interviews can be enriched by considering digital tools as purely heuristic instruments, not instruments that connect directly to their existing ways of interpretation of data, but that should be considered at the very beginning of the research process, when one is still considering what collections to reuse, or how the characteristics of one's own data relate to open source data. Besides raising these epistemological concerns, the workshop participants pointed to some more mundane concerns. Explanation of linguistic approaches would be better appreciated in more lay terms, following a step-by-step pathway from meta to concrete level, following the way in which the introduction of linguistic tools was prepared for the workshop. Participants, many of whom admitted to a lack of technical proficiency, felt that dejargonising the approaches, and preparing a simplified 'layer of information' would make the tools more appealing to unfamiliar users and make getting started easier.

Third, users are also concerned about where data are going when online tools are used for processing their interview data, prompting thinking about how these tools might be better aware of legal and ethical issues, for example adding explicit GDPR-compliant data processing agreements to allay worries. Finally, users pointed to a great need for contextual information about, and metadata for, the data collection and processing, when interview data sources are used. Inviting language resources and scholars across languages and different disciplines certainly enriched the meeting experience.

Corti (corti@essex.ac.uk), UK Data Archive and Silvia Calamai (silvia.calamai@unisi.it), Universita di Siena, sede Arezzo and Norah Karrouche (karrouche@eshcc.eur.nl), Erasmus Universiteit Rotterdam, The Netherlands and Jeannine Beeken (jeannine.beeken@essex.ac.uk), UK Data Archive and Christoph Daxler (draxler@phonetik.uni-muenchen.de), Ludwig-Maximilians-Universität München and Henk van den Heuvel (H.vandenHeuvel@let.ru.nl), Radboud Universiteit Nijmegen
