

Croatian Memories

Speech, Meaning and Emotions in a Collection of Interviews on Experiences of War and Trauma

Franciska de Jong^{1,2}, Arjan van Hessen², Tanja Petrovic³, Stef Scagliola¹

¹ Erasmus Studio, Erasmus Universiteit Rotterdam, The Netherlands

scagliola@eshcc.eur.nl

² Human Media Interaction Group, Universiteit Twente, The Netherlands

f.m.g.dejong, a.j.vanhessen@utwente.nl

³ Documenta Center for Dealing with the Past, Zagreb, Croatia

tanja.petrovic@documenta.hr

Abstract

In this contribution we describe a collection of approximately 400 video interviews recorded in the context of the project Croatian Memories (CroMe) with the objective of documenting personal war-related experiences. The value of this type of sources is threefold: they contain information that is missing in written sources, they can contribute to the process of reconciliation, and they provide a basis for reuse of data in disciplines with an interest in narrative data. The CroMe collection is not primarily designed as a linguistic corpus, but is the result of an archival effort to collect so-called oral history data. For researchers in the fields of natural language processing and speech analysis this type of life-stories may function as an *objet trouvé* containing real-life language data that can prove to be useful for the purpose of modelling specific aspects of human expression and communication.

Keywords: multimodal spoken word corpus, narrative data, oral history, data reuse

1. Introduction

The 20th century has been characterized as the ‘The Era of the Witness’ (Wieviorka, 1998), while the 21st century has been labelled as ‘The End of Forgetting’ (Rosen, 2010). The latter concept refers to the wide availability of technology for bearing witness and preserving memories. But as the project **Croatian Memories** (CroMe)¹ illustrates, technology has more to offer, and can play a crucial role in programmes that facilitate reconciliation in post-conflict regions. The project’s goal was to collect personal testimonies on war and trauma from population groups that are hitherto underrepresented in the Croatian public realm and to arrange access to them through an online platform. The result is a so-called oral history interview collection in which citizens from all social layers and regions in Croatia reflect on three major timeframes:

- WWII
- the period of socialist Yugoslav
- the war of the nineties.

The length of the interviews varies between 40 and 200 minutes and their structure is determined by a biographical semi-structured questionnaire. The video-recorded interviews plus the manually generated annotation layers are indexed at fragment-level, and have all been transcribed in Croatian and translated into English, yielding the basis for

searchable time-aligned subtitles in both languages.

The metadata includes information related to the interviewees (e.g., profession, religion, age), and to the interviews, such as summaries, transcripts, and details on when and where the interview was conducted.

The interview collection and some basic metadata is made public through an open access streaming video platform² with rich search functionalities, launched in the fall of 2013 and hosted by the human rights organisation Documenta (Zagreb)³. At present all interviews have been published as open data, in accordance with the permission granted by the narrators. However, in order to constrain access to any passage or metadata category that may prove to jeopardize the privacy of respondents, there is the possibility to seal off sensitive passages and/or to install embargos.

For scholarly use and under strict conditions only, there is an alternative access option that offers access to the entire collection and to all metadata, including any restricted parts, in accordance to the conditions agreed with the narrators. The catalogue for this scholarly access platform will be hosted by DANS, the academic archiving services institute in The Netherlands.⁴

A next development step envisaged is the enhancement of the current infrastructure with tools that encourage and facilitate both reuse and secondary analysis of the

¹ Cf. postyugoslavvoices.org/?page_id=23

² Cf. www.croatianmemories.org/

³ Cf. www.documenta.hr/en

⁴ Cf. www.dans.knaw.nl

collection. Among the options considered is the creation of a virtual research environment that provides access to the interview data as well as to a set of online content analysis tools. Such a service-model approach would allow the processing of data without the need of having to download the interviews and/or transcriptions, and other annotation layers.

This paper aims to present the content of the collection and its characteristics (Section 3 and Section 4), as well as a vision on how the CroMe collection offers a basis for a multidisciplinary research agenda (Section 5), covering the perspectives of a wide range of areas, including speech processing, linguistics, text mining, memory studies, history and psychology. In order to support potential users in deciding whether the materials can serve their needs, overviews are presented with regard to the characteristics of the interviewees and their distribution over the collection. In Section 2 we will first present our vision on how digital oral history data can be reused for purposes unrelated to the primary aim of creating a specific oral history collection.

2. Oral History: Collection *versus* Corpus

Although the term ‘oral history’ has various meanings, there is common understanding that its purpose is to create oral accounts on personal history in an interview setting. A distinction can be made according to the aim of the effort of collecting interviews. This can be either answering a specific research question, or documenting people’s experiences as an archival effort with future listeners in mind (Freund, 2009).

The so-called ‘digital turn’⁵ that set in at the end of the 20th century with the increasing availability of easy-to-use recording devices and ease-of-access to digital data, has increased the attention for oral history data outside the field of history. Innovations from the field of information retrieval and natural language processing provide the tools for the (semi-)automatic indexing of interviews at all layers, including the fragment-level. Moreover, reuse and sharing of interview datasets is more and more advocated and the emerging availability of infrastructures that can be deployed for both digital born and digitized oral history collections has become a strong incentive for setting up novel research agendas.

CroMe is the result of an archival effort. For researchers in the fields of natural language processing and speech analysis such collections may function as an *objet trouvé* that provides real-life language data that can prove to be useful for the purpose of modelling specific aspects of human expression and communication. Rather than a corpus with a specific-purpose design, it is a collection of narratives with a variety of possibilities for scholarly use.

3. Collection Design

The collection shares a number of characteristics with other

digital video oral history collections, such as the project *Forced Labor 1939-1945*⁶ and the *Visual History Archive*, the online portal created by the USC Shoah Foundation⁷. Croatian Memories distinguishes itself through the combination of scale (>400 interviews), the principle of open access/open data, the multiple options of the platform for exploring and searching the content at fragment-level and the multilingual dimension, but foremost by its commitment to connect to innovative research practices across disciplines and to support the potential for reuse of the audio-visual interviews. The extent to which the platform content is suited for reuse may of course vary across disciplines. In principle there was a limit to the types of reuse that could be anticipated at the time of designing the protocol for data collection and annotation. Therefore it may well turn out that the collection is not meeting the full set of requirements for certain fields, and that additional investments in annotation would be needed. Also, as explained in Kemman et al (2014), interface design for scholarly use of history collections requires careful interaction with representatives of the envisaged user groups.

3.1 Recording guidelines and interview protocol

The interview sessions were conducted according to guidelines, both for the method of interviewing and the use of recording equipment.

Recording protocol

The recording protocol was adjusted to the requirement of making the interviewees feel as comfortable as possible. Most of the interviews were recorded in the interviewees’ homes; some were conducted at the premises of local community service organizations or at places in which the interview team resided during their field trips. The A/V recording quality (for aspects such as sound and vision) that could be realised was dependent on these local conditions. This imposed some limitations for capturing the non-verbal part of the speech: body posture, hand gesture, head gesture, but ‘tone of voice’, prosody, and expressions such as laughter, hesitations and sighs are a well-covered aspect in the recorded material. Therefore the data is suited as a basis for human behaviour analysis as pursued in fields known as social signal processing and/or affective computing.

The rationale for the recording protocol adopted was that the eventual recorded material should in any case allow researchers to explore and detect details from the sound tracks demonstrating the variety in aspects such as tone, volume, rhythm of speech, etc. Presumably the speech signal also contains traces of the emotions that are expressed, such as grief, anger or resignation. Also worthwhile investigating is the seemingly lack of emotions that can be observed in the narratives of traumatized people. As the study of these semantic layers in narrative speech data is still in its infancy, insights in the optimal recording conditions were only partially available. Some obvious

⁵ <http://www.thedigitalturn.co.uk/>

⁶ Cf. www.zwangsarbeit-archiv.de/en/, and also Plato (2010).

⁷ Cf. <http://sfi.usc.edu/>

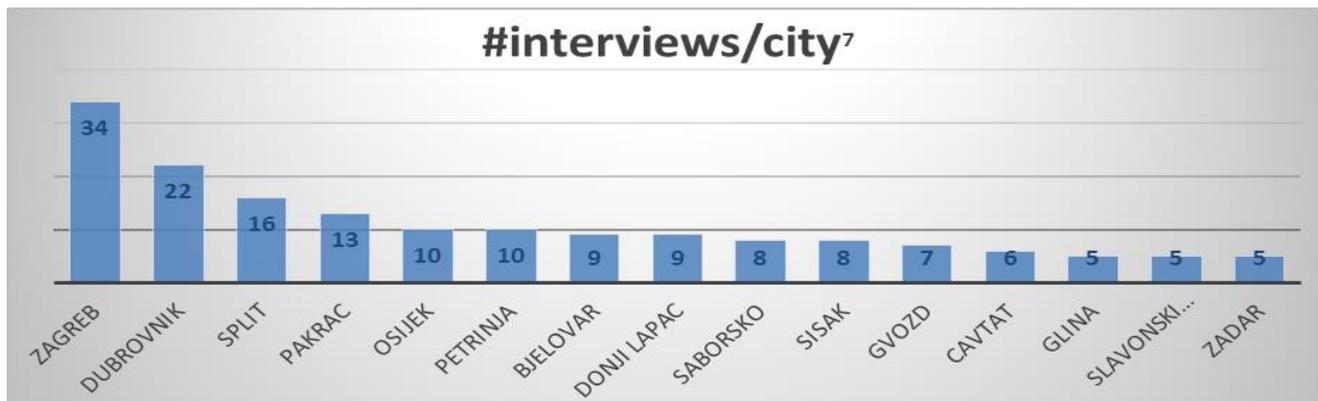


Figure 1: Distribution of “Cities” where the interviews were recorded.

research desiderata, such as the use of close camera and microphones, or even more invasive equipment such as sensors for measuring bodily effects of emotion, were incompatible with the primary goal of the collection: giving voice to people from minority groups.

Interview methodology

The interview methodology adopted was tuned to the creation of process-generated oral history data. As explained in Section 2, the CroMe collection is the result of an archival effort to document experiences of war and trauma in a balanced way by interviewing a considerable number of people on the basis of a range of similar questions, the so-called topic list. No specific research question had to be answered, the interviews recorded are semi-structured and apart from giving voice to a specific group of citizens, their content is meant to offer suitable material for scholarly reuse and comparative analysis.

In principle the topic list followed a biographical/chronological order, but as mentioned, the interview team often had to adapt their approach to suit the needs of the narrators. In order to create a flow of speech, sometimes traumatizing events had to be narrated first, before the interviewee could provide the necessary context on his or her biographical background that is needed for a future viewer to understand the logic of the story.

3.2 Facts and Figures on Interviews and Interviewees

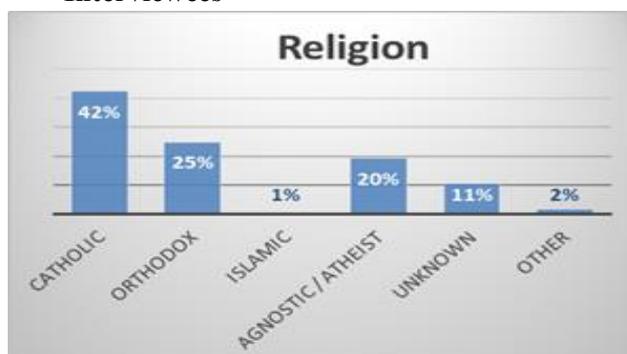


Figure 2: Distribution of “Religion” amongst the interviewees

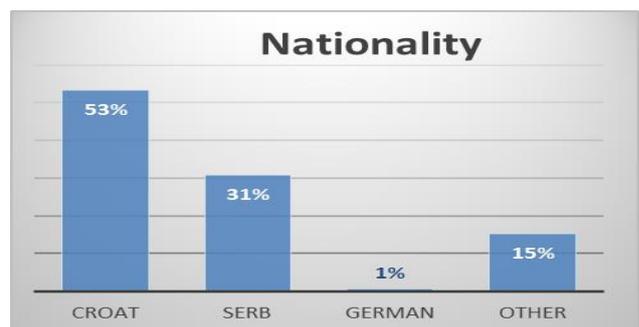


Figure 3: Distribution of “Nationality” amongst the interviewees

Care was taken to balance the background characteristics of the interviewees. With regard to gender, around 63% of the interviewees is male *versus* 37% female. There is a good geographical spread of the interviewees’ place of residency, which in most cases is also the place where the interview was conducted. This is partly thanks to Documenta’s broad network within Croatian and Serbian communities of victims all around the country, including those parts that before 1995 were mostly populated by

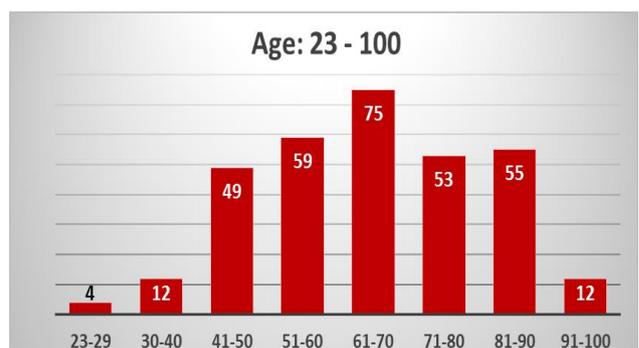


Figure 4: Distribution of “Age” amongst the interviewees.

Serbs. As depicted in Figure 1 the interviews are from over 126 unique places and the highest number of interviews conducted in the same place is 34 (for Zagreb).

The figures 2-5 in this paper give an indication of the spread across various other dimensions.⁸ It should be noted however that in this type of collection one can only strive

⁸ The figures hold for the 328 interviewees processed by May 2013, i.e. around 80% of the collection. A record of future

changes in the facts and figures will be kept at this webpage: http://postyugoslavvoices.org/?page_id=1317

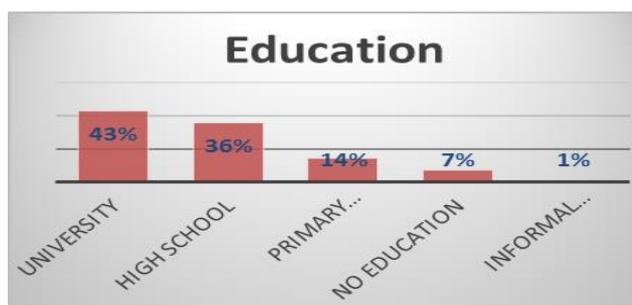


Figure 5: Distribution of “Education” amongst the interviewees.

for representativeness from a social-geographic perspective. The CroMe collection does not meet the requirements for balance typically adopted in the development of corpora that are used for deriving speech models, nor those for the selection of population samples commonly applied for quantitative studies in the social sciences. In this case the collection is dominated by narrators that have a so-called 'urge to tell'. This points to the fact that other people, engaged in similar circumstances, may speak about their experiences in quite a different way, or not at all.

4. Quality Control and Validation

In the creation of the collection, measures have been taken in order to guarantee that (i) the interviews meet the commonly acknowledged criteria for good oral history practice and (ii) the annotation layers generated are complete and consistent with regards to the adopted metadata model and other standards for quality. To this end, a number of validation steps have been performed which will be described in Section 4.2.

4.1 Quality Criteria

4.1.1 Oral history quality criteria

The protocols for conducting an oral history project have been topic of concern and debate ever since it has become an established academic field. The widely shared insights in the importance of guidelines for interviewers, the protection of the interest of the interviewees both during the interviews and in the handling of any agreed constraints for access and use, as well as transparency on the measures and policies adopted, have been incorporated into the interview methodology and the portal design. Inspiration was derived from the framework of value-sensitive design (Friedman *et al.*, 2006; Van den Berg *et al.*, 2010).

The following essential elements have been covered:

- Setting up guidelines regarding the interview process (preparation, recording- and interviewing techniques, topic list, post-interview protocol)
- Setting up guidelines for creating metadata of the interview
- Training, monitoring and providing feedback to interviewees and metadata creators

- Arranging an informed consent protocol tailored to the exigencies of the project
- Designing a dedicated metadata model, covering:
 - o Contextual information at collection level (topic list, method of recruitment and of informing interviewees, transcription policy (readability *versus* exactness of paraphrase, handling translation problems))
 - o Annotations at the interview-level (date/place interview, background variables interviewer/interviewee, summary, etc.).
 - o Annotations at the fragment-level (key words, topics labels, ...).
 - o Personal details of the interviewees. (Note that contact details of narrators are not included in the dataset, and are only known to the curators of the collection.)

4.1.2 Recording quality criteria

As indicated above, the recording process often had to be adjusted to the interviewee’s needs. In order to optimize the A/V-quality, the recordings were regularly checked, and feedback was given to recording team concerning composition, use of zoom and sound. All A/V-recordings (incl. the audio) were made with a high quality camera. For the workflow, and the final presentation on the website the resulting HD-videos (>20GB/interview) were transcoded into a MP4-version (≈500MB/interview) with an audio quality of 48kHz (mono). The HD-versions are stored at DANS for eventual future TV-broadcasts, the MP4-versions are accessible via the website of Croatian Memories. These specifications are known to be adequate for the purposes indicated (broadcasting and streaming, respectively), and in line with the standards advocated by DANS⁹.

4.1.3 Quality criteria for metadata and annotation

As mentioned in Section 4.1.1 it is crucial for the reuse of the data for oral history research that an adequate metadata model is adopted. The practice until recently in the field of oral history is that each project draws on previous similar projects known within national boundaries or a specific network. Collections created within libraries or museums will usually adhere to the Dublin Core¹⁰, whereas civil society organisations tend to create their own systems and schemata. The experience in CroMe and the sibling project BiHMe¹¹ have contributed to the development of a standard for multilingual video oral history archives, that is flexible and makes initiators of a project conscious of choices they can make, and of how these will determine their potential audience and the societal and academic valorisation of the new collection. Stimulating this awareness connects to the requirements for innovative use of ICT in research, such as interoperability, sharing and the reuse of data.

In CroMe additional annotation layers have been created in

⁹ <http://www.dans.knaw.nl/en/content/audio-and-visual-data>

¹⁰ Dublin Core Metadata Initiative: <http://dublincore.org>

¹¹ <http://www.bosnianmemories.org>

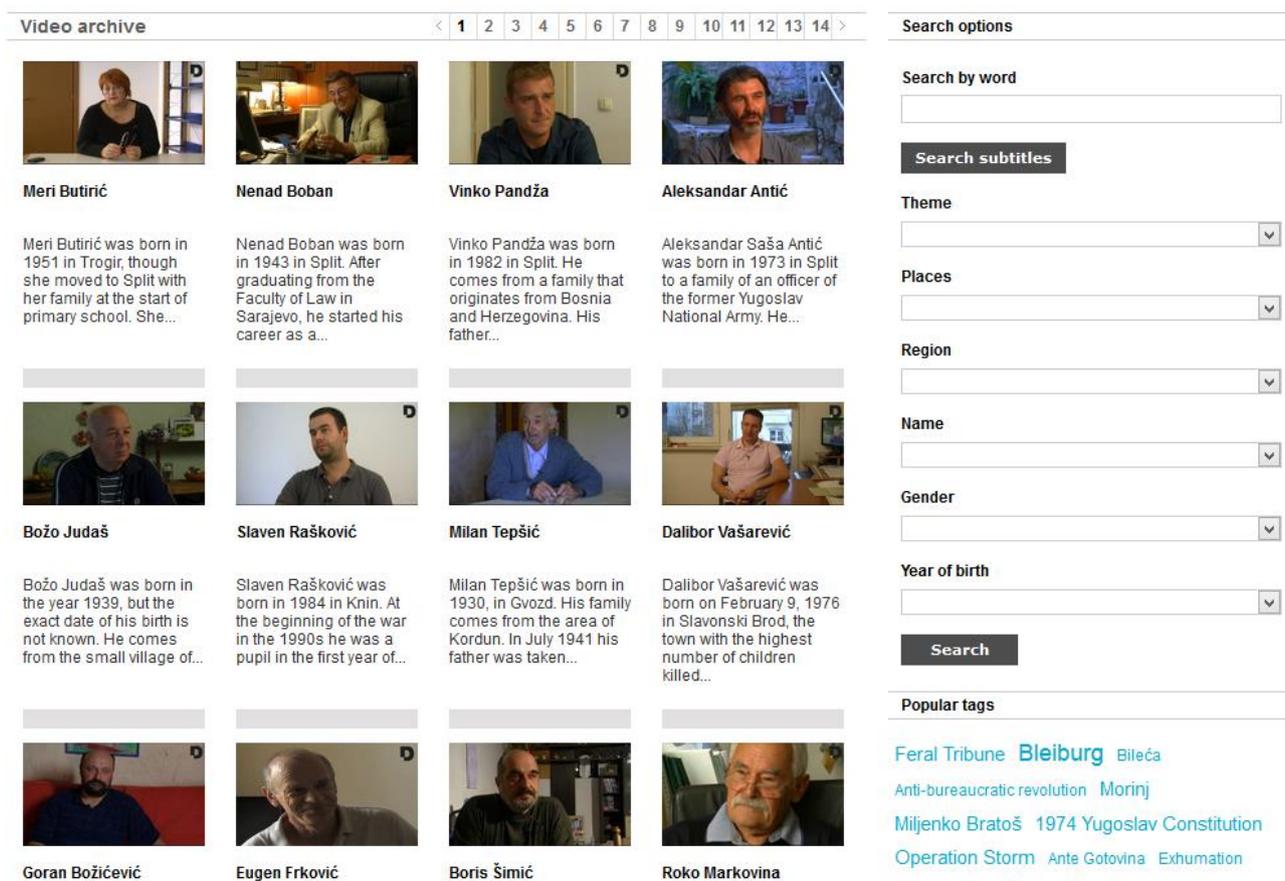


Figure 6: The search interface for the CroMe-video archive. One can search in the “properties” of the interviewee (name, gender and year of birth), in the time-dependent metadata (themes, places and regions/locations mentioned in the interview) and full text in both the Croatian transcriptions and English translations. The search results in a number of clickable thumbnails. Clicking results in playing of the interview fragment starting 1 second before the occurrence of the search term.

both Croatian and English, in order to facilitate the online search and presentation of the collection for an international audience.

Figure 6 illustrates the two modes of searching that are enabled by the annotation layers that have been generated: (i) full text search, based on transcripts (both for the original language, and for the translations in English), and (ii) faceted search, based on the annotations. Some layers come with timestamps, which enables the search engine to return a result list with links to the relevant fragments of the recordings. In order to meet the criteria for adequate subtitling, certain crucial choices had to be made regarding the length and the level of ‘reduction’. The aim has been to select what would be adequate for presentation on the screen and still would yield a sound basis for content analysis. Sometimes the creation of transcripts and adequate translations was challenging, for example in the case of interview passages in which interviewees preferred to move around or spoke incoherently, or in language or dialect that is difficult to understand anyhow.

For most interviews the overall amount of reduction in the Croatian subtitles could be kept to a minimum: ‘near to exact’ for the words used, including repetition, hesitations, etc. but mostly without the ‘ehs’ and ‘ehms’. For the

English subtitles a more principled choice had to be made to keep a proper balance between informativeness and readability, as in principle loss in translation cannot be avoided. Also because of the limited availability of resources (volunteers fluent in Croatian and English), it was decided to give priority to accessibility and readability.

In some cases annotations have been added to the translation to explain the choices that have been made. This type of metadata is available for research purposes. Crucial, and also rather exceptional for this type of collections, is the availability of subtitles that have been translated *and* indexed. Whatever the type of envisaged use may be, it offers the possibility to detect the availability of relevant data. Once relevant material has been identified, the original content can be scrutinized more in-depth, which may lead to the decision to arrange more detailed translations of specific (parts of) interviews.

4.2 Control and Validation

Specific control and validation steps were carried out for the following content processing steps: transcription, translation, subtitling and annotation. They will be summarized in the following subsections.



Goran Božičević

Goran Božičević was born in 1962 in Zadar. His father was a highly positioned officer in the Yugoslav National Army. When he was in his fourth year of high school his family moved to Zagreb where Goran started studying physics at the Faculty of

Figure 7: The result of searching the subtitles (both transcriptions and translations) for the query “The Hague”. This term appears twice in the interview with Goran Božičević. Clicking on a subtitle in the search result (*in this case two*) causes the player to jump to the place in the interview where the word is used.

4.2.1 Transcription and subtitling

The transcription and translation of the interviews, and its conversion into subtitles was done by various people in different countries. The handling of Croatian transcriptions was coordinated by Documenta; the edit tool used was Open Source application SubtitleEdit¹². The resulting UTF-8 transcriptions were first checked and adjusted manually by the University of Twente, then outsourced for translation by native-Croatian language students, and finally adjusted to the requirements of subtitling. In accordance with the “best practise guidelines” of broadcast companies¹³, it was decided that the maximum length of each subtitle should not exceed 100 characters (including spaces) and that no more than 2 lines of subtitles should be shown on the screen. An application was made that splitted longer phrase in parts of max. 100 characters. To maintain the parallel between the English and the Croatian version, splitting was done for both languages. As a result, the amount of English and Croatian subtitle lines is equal for each interview. In general, subtitles are heavily edited versions of what is said. However, because of the envisaged future use by researchers it was decided for Croatian to stay as close as possible to the original wordings. For certain fragments, this may require some fast reading.

4.2.2 Translation

To address the well-known problem of translating non-standard and/or incoherent language into a standard version of another language (Hooks, 1995), an application was made that compared the Croatian and English subtitles for

length. Over all files, the amount of English words/file was slightly higher (10%) than the amount of Croatian words/file (9885 *versus* 8969). If the difference per subtitle line exceeded 20%, the line was marked and sent to a native speaker for an additional check. Apart from these cases, roughly 15% of the transcriptions and translations were randomly checked by a third person.

4.2.3 Annotation

The interview transcripts were the basis for annotation, plus a 25 fields metadata scheme, including: personal information on the interviewee (e.g., name, address, education, religion, nationality), status of the interviewee (e.g. civilian, member of the army, refugee, internally displaced, disabled, family member of killed/missing person, imprisoned, in hiding, member of the resistance), information on the interview (e.g. place and date of the interview, interviewer, videographer, summary of the interview), place/region and period of remembrance, and keywords (up to three per interview selected from a list of 100). For a pilot collection of 50 interviews the usefulness and quality of annotations was assessed during an international multidisciplinary workshop¹⁴ (December 2012, Rotterdam).

5. Next Stage: Development of Research Agendas

Besides the primary goal of stimulating the process of reconciliation activities in Croatia, the CroMe project intends to draw scholarly attention to oral history data. Currently, interview collections of this type are seldom

¹² Cf. www.nikse.dk/subtitleedit

¹³Cf. www.bbc.co.uk/guidelines/futuremedia/accessibility/subtitling_guides/online_sub_editorial_guidelines_vs1_1.pdf

¹⁴ thehagueinstituteforglobaljustice.org/index.php?page=Events-Events-Events-Balkan_Memories_expert_meeting&pid=166&id=79

reused by researchers who are not involved in the creation of the collection, while rich annotation and state-of-the-art digital search and navigation interfaces within the personal narratives offer ample possibilities for research in several disciplines. The recurring themes addressed in the topic list, create a fruitful potential for comparative research. A concrete step as a basis for comparison is the creation of a (smaller) sibling collection in Bosnia and Herzegovina, that was launched in November 2013.¹⁵

A whole range of analysis tools rooted in the field of text and speech processing are available for take-up for audio-visual datasets. There is ample evidence that this type of research could bring new insights into the use of language and in the feasibility of automatic metadata generation and of larger-scale content explorations within and beyond collection boundaries (De Jong *et al.*, 2008; Oard *et al.*, 2002). Moreover, in a collaborative project combining insights from psychology and speech analysis, a study has been conducted based on a pilot set of 50 CroMe video interviews that shows the potential of this type of data for comparing verbal patterns with non-verbal features (Truong *et al.*, 2013).

There is some insight in the diversity in the way stories are interpreted and analysed across fields (Hyvärinen, 2012), but the study of the implications of personal accounts and their potential for studying human communication is still in its infancy. Multidisciplinary programmes involving e.g. linguistics, memory studies, history, psychology, media studies and the study of transitional justice, could help to increase the value of personal narratives for the understanding of how people construct meaning and identity by telling stories and giving testimony through the agency of memory and language. The multiple layers of meaning offered by the CroMe collection could prove to be of great added value for the maturing of collaboration between humanities, social sciences and technology.

Acknowledgements

The project Croatian Memories (CroMe) was funded by the Ministry of Foreign Affairs of the Netherlands.

References

- Freund, A., (2009). Oral History as Process-generated data. *Historical Social Research*, 34 (1): 22-48.
- Friedman, B., P. H. K. Jr., and A. Borning, (2006). Value-Sensitive Design and Information Systems. In: *Human-Computer Interaction in Management Information Systems: Foundations*. (chapter 16). M.E.Sharpe, Inc.
- Hooks, B. (1995). "This is the Oppressor's Language/Yet I Need it to Talk to You": Language, a Place of Struggle. In: *Between Languages and Cultures: Translation and Cross-Cultural Texts*, ed. by Anuradha Dingwaney and Carol Maier. Pittsburgh, PA: University of Pittsburgh Press. 295-301.
- Hyvärinen, M., (2012). Prototypes, Genres, and Concepts: Travelling with Narratives, In: *Narrative Works*, 2: 10-32
- Jong, F.M.G. de D.W. Oard, W.F.L. Heeren, and R.J.F. Ordelman, (2008) *Access to recorded interviews: A research agenda*. In: ACM Journal on Computing and Cultural Heritage (JOCCH), 1 (1). 3:1-3:27.
- Kemman, M., S. Scagliola, F.M.G. de Jong, R.J.F. Ordelman (2014), *Oral History Today – Exploring Oral History Collections*. To appear in: Proceedings Conference Digital Humanities Benelux 2014.
- Oard, D.W., et al (2002). Cross-language access to recorded speech in the MALACH project. In *Proceedings of the Text, Speech, and Dialog Workshop*. Brno, Czech Republic. 197–212.
- Plato, A. von, (2010). Reports from Germany on forced and slave labour. In A. V. Plato, A. Leh, & C. Thonfeld (eds.), *Hitler's slaves: Life stories of forced labourers in Nazi-occupied Europe* (pp. 23-36). Oxford/New York: Berghahn.
- Rosen, J. (2010), The Web Means the End of Forgetting. In: *New York Times*, August 2010.
- Truong, K., et al., (2013), Emotional expression in oral history narratives: Comparing results of automated verbal and nonverbal analyses. In: *Proceedings of CMN2013*, Hamburg.
- Van den Berg, H, S. Scagliola, F.Wester (eds.), (2010), *Wat veteranen vertellen: verschillende perspectieven op biografische interviews over ervaringen tijdens militaire operaties*. Amsterdam University Press. See also: <http://www.watveteranenvertellen.nl>
- Wieviorka, A., (1998), *L'Ere du Temoin*, Paris.

¹⁵ For details on the BiHMe collection, cf. postyugoslavvoices.org/?page_id=26