# Ravensbrück Interviews: How to Curate Legacy Data to Make it CLARIN Compliant

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

**Stefania Scagliola**
Independent Researcher
Rotterdam, The Netherlands
scagliolas@gmail.com

**Fabio Ardolino**
Università di Siena
Siena, Italy
fabio.ardolino@unisi.it

**Christoph Draxler**
Ludwig Maximilian University
Munich, Germany
draxler@phonetik.uni-muenchen.de

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

**Henk van den Heuvel**
Radboud University
Nijmegen, The Netherlands
henk.vandenheuvel@ru.nl

## Abstract

This paper describes the preparatory phase of a CLARIN-funded project called 'Voices from Ravensbrück', which aims to introduce a new type of corpus in the CLARIN resource family called 'Oral Histories'. The first task consisted in curating and transcribing a set of interviews conducted by the Italian author A.M. Bruzzone with five Italian survivors of the Ravensbrück concentration camp back in 1977. This posed considerable challenges inherent in integrating legacy data from the pre-digital era in the CLARIN infrastructure. The second task was exploring the potential of automatic speech transcription for this type of oral history data. The third element of this exploratory phase was identifying potential partners and suitable data for creating a multilingual collection of existing oral history interviews with survivors of concentration camp Ravensbrück. These preparatory steps were necessary to move to the final phase of our project and realise our overall objective of creating a resource family compliant with CLARIN standards and enabling scholars to analyse interviews from a comparative multilingual and multidisciplinary perspective

## 1    Italian Interviews: Curation

In 1976, Anna Maria Bruzzone and Lidia Beccaria Rolfi collected testimonies of 5 Italian women who had been deported to the Nazi concentration camp Ravensbrück. The analogue archive containing the Italian interviews was donated to Siena University by Anna Maria Bruzzone's niece. Siena University digitised all the recordings of Bruzzone's archive according to IASA standards (.wav format, 96000 Hz, 24 bit). Preservation copies were created and were structured as follows: *i*) audio files, *ii*) photos of the carrier, *iii*) metadata.

   Bruzzone's Ravensbrück collection consists of 14 audio cassettes, with a total duration of about 18 hours and 20 minutes, and contains four long interviews. We know that Anna Maria Bruzzone transcribed the recordings step by step for her publication, writing everything down that she heard. Still, unfortunately, the handwritten transcriptions were lost. Her book, titled *Le donne di Ravensbrück* (Beccaria Rolfi, Bruzzone, 1978, ed. 2021; Figure 1), was divided into 4 sections, each one related to a deportee testimonial (except for the last): Lidia Beccaria Rolfi, Bianca Paganini Mori, Lidia Borsi Rossi, the sisters Nella Baroncini Poli and Lina Baroncini Roveri. In 2016, the book was translated into German (*Als Italienerin in Ravensbrück. Politische Gefangene berichten über ihre Deportation und ihre Haft im Frauen-Konzentrationslager*; Beccaria Rolfi, Bruzzone, 2016). In order to re-use the oral archive and offer insights into valuable legacy data, it was necessary to first define a proper legal framework for

---

Figure 1: Cover of the volume curated by Lidia Beccaria Rolfi and Anna Maria Bruzzone (2021 edition).

issues related to copyright and privacy. According to Italian law, the copyright of the interviews has been passed on to Bruzzone's niece Paola Chiama and to the heirs of Lidia Beccaria Rolfi, co-author of the volume. In the act of donation to Siena University, P. Chiama has authorised the re-use of the archive for research, dissemination, and teaching purposes (after giving her appropriate advance notice). As for the privacy issues, the legal framework for the re-use and the release (via a CLARIN repository) of the digitized Italian interviews has been accurately defined with the advice of Giuseppe Versaci, lawyer and data protection officer of Siena University and member of the CLARIN Legal and Ethical Issues Committee. In detail, it was considered that, according to Italian law, access to the audio interviews is in any case possible on the basis of the legitimate interest of research forty years after the recordings were made – or sixty years, in case of documents containing data that is likely to reveal the state of health, sexual life or confidential family relationships (Legislative Decree 22/01/2004, art. 122). This normative reference is coupled with art. 2 Legislative Decree 196/2003 (*Italian privacy code*), which gives the heirs the possibility to exercise rights concerning deceased persons ("The rights referred to in Articles 15 to 22 of the Regulation concerning the personal data of deceased persons may be exercised by any entity having a vested interest or acting to protect the data subject as the latter's agent, or else on household-related grounds deserving protection").

Notwithstanding this legislative backing, for ethical reasons, it was decided to inform all legal heirs of the five interviewees about the initiated project to obtain their additional consent. With this aim, two distinct documents have been prepared: a) a private letter containing all the information about the project's backgrounds and goals, and b) a document with detailed information on how the data was processed. The final texts of these documents were submitted to the data protection officer to verify the full compatibility with the Italian and EU legal frameworks.

Once the legal provisions of the project were clear, the heirs of the interviewees had to be traced back. In this phase the support of two Italian associations, the ISR, *Istituto Spezzino per la storia della Resistenza e dell'Età Contemporanea*, and the ANED, *Associazione Nazionale Ex Deportati* was crucial. For all the interviewees, a direct living heir was identified and contacted. Four of the contacted heirs (Aldo Rolfi, Eligio Roveri, Giorgia Poli, and Anna Maria Mori; the last also on behalf of her sister Paola Mori) have notified their full consent to the project, while an answer is still awaited from Borsi Rossi's heir. Each heir has also received a digital copy of the interview related to his/her relative. Though not set in legal obligations, this approach reflects the type of relationship that oral historians have with their narrators and their next of kin, one in which 'a shared authority' on the output of the research is a common practice (Frisch, 1990).

During the project, metadata files compliant with CLARIN standards were created. They were based on an existing CMDI profile developed for an earlier CLARIN interview data curation project named *Oral History Interview* (it can be found as [OralHistoryInterview](#)[1]). However, the profile was created for born-digital interview data generated through a new project. Whilst it is true that this profile provides a set of components useful to report pertinent traits of oral interviews (e.g., interviewee and interviewer specifications, interview methods, audio characteristics, and annotation protocols), it only partially adapts to speech materials like the ones contained in the Anna Maria Bruzzone archive: indeed, legacy data pose peculiar challenges on the basis of the relationship among the original documents (analogue carriers), their digitised versions, the documental units (i.e., a single interview which might be contained in more than one single analogue carrier). Therefore, the specificities of Bruzzone's archive with its original analogue recordings require a partial reorganization of the components of the metadata file, to abide to the archival principle of provenance. The following components must be added: *a*) information about the context in which the interviews were conducted; *b*) information about the process of digitisation of the interviews.

A rearrangement of the original CMDI OH profile component is necessary to meet these new requirements. This could be the addition of two components – *Context of Creation* and *Context of Digitisation* – or the addition of one field with the name *Description* (in the example of the scheme shown in the appendix, the two specific fields have been added). For entering the resource family in the online catalogue of CLARIN, it is important to take into account that three levels of access are required: 1) to the entire collection, 2) to the subcollections in a particular language, and 3) to the single interview. To this end, the [lat-corpus profile](#) of the TLA can be used.

## 2    Transcription

Under the CLARIN umbrella, starting from the 2016 CLARIN Oral History workshop in Oxford, a group of experts interested in speech data with very different backgrounds – oral history, computational linguistics, anthropology, sociolinguistics, phonetics, and phonology – started exploring how technology can be integrated into research that involves spoken narratives contained in oral archives, thus creating a network called *[Speech data and Technology](#)*[2] (Draxler et al., 2020; Scagliola et al., 2020). For making the initial transcriptions, the T-Chain was used. The T-Chain is an open-source transcription workflow for interviews that can be accessed via the [TranscriptionPortal](#)[3], where scholars can upload audio files, select the spoken language and, for some languages, a Language Model, and process their files. The digitized version of the five interviews was used as a "stress test" to ascertain the T-Chain's potential with legacy data. The T-Chain proved to be extremely useful for phonetic and word alignment. At the same time, its use in this preparatory project outlined the workflow's technical, economic and organizational limits. If one can rely on a previous, accurate transcription, the alignment (both at phonemes and words levels) appears to be rather good. Figure 2 shows a graphical representation of an excerpt from an automatically generated word alignment.

For this alignment, the original speech signal was processed using the Google automatic speech recognizer for the Italian language, which returned an orthographic transcript. This transcript was corrected manually, and then processed with the MAUS web service (Kisler et al., 2012) in order to obtain the alignment. The final result is a three-tiered transcription, which specifies the position of the orthographic words, their expected canonical pronunciations, and the sounds of the actual utterance: the actual utterance is often quite different from the expected pronunciation, due to coarticulation, accented speech, and low signal quality. This position is generally given either in timestamps or sample points relative to the start of the utterance. Table 1 shows the same excerpt as a table, ready for import into a spreadsheet, database system, or statistics package.

---

[1] [https://catalog.clarin.eu/ds/ComponentRegistry/#/; look for the profile OralHistoryInterview](https://catalog.clarin.eu/ds/ComponentRegistry/#/)
[2] [https://speechandtech.eu/](https://speechandtech.eu/)
[3] [https://speechandtech.eu/oh-portal](https://speechandtech.eu/oh-portal)
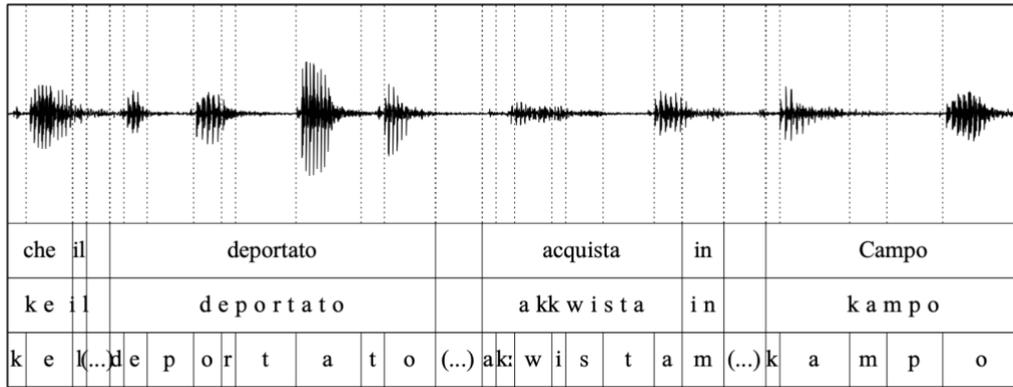
Figure 2: Excerpt of an automatic alignment: the top shows the raw audio signal waveform, the three tiers display the orthographic words, their expected pronunciation, and the actual pronunciation, each with its exact position within the signal.

| Line | Begin | Tier | Text |
|------|-------|------|------|
| | | ... | |
| 47 | 21.6800 | ORT-MAU | che |
| 48 | 21.8200 | ORT-MAU | il |
| 49 | 21.8500 | ORT-MAU | |
| 50 | 21.9000 | ORT-MAU | deportato |
| 51 | 22.6000 | ORT-MAU | |
| 52 | 22.7000 | ORT-MAU | acquista |
| 53 | 23.1300 | ORT-MAU | in |
| 54 | 23.2200 | ORT-MAU | |
| 55 | 23.3100 | ORT-MAU | Campo |
| | | ... | |

Table 1: Table rendering of the time-aligned orthographic transcription on the word level in Figure1.

The audio files originally come in archive quality (96 kHz sample rate, 24-bit linear quantisation, stereo) to result in a data rate of 0.58 MB/s. Many speech processing tools do not require such high data rates, and thus the audio files were downsampled to 16 kHz, 16-bit linear quantisation, mono for a data rate of 0.032 MB/s (a >90% reduction). This downsampling is done before the audio data enters the T-Chain to reduce the amount of data to transfer.

After this pre-processing step, we verified the potential of the ASR (*Automatic Speech Recognition*) system for the Ravensbrück interview data. Since we cannot rely on academic open-source ASR software for the Italian language, we were forced to use a commercial system, which obviously was not trained on legacy data. Given the varying acoustic quality of legacy data, the variable interview settings, and the presence/absence of vernacular speech throughout the interviews, we decided to proceed with a qualitative/quantitative evaluation of the effective ASR performances.

The ASR outputs were evaluated for two example audio documents: *i*) the initial part (48 minutes, of which 44.55 spoken) of the interview with Lidia Beccaria Rolfi (file name: BRZTO061a); *ii*) the initial part (31 minutes, of which 29.30 spoken) of the interview with Lina Baroncini Roveri and Nella Baroncini Poli (file name: BRZTO067a).

The two samples have been selected since they represent a pair of opposites: the interview with Lidia appears to be linear, with a sharp prevalence of one of the actors (Lidia), while the interview with Lina and Nella is rather complex, presenting a more intricate diarization among four speakers (Anna Maria Bruzzone, Lidia, Lina, and Nella) and a more vernacular style.

The two ASR outputs for each file have been compared with transcriptions of the same samples that were manually corrected to obtain the highest achievable accuracy. In line with the most widespread standards, a threshold of 30% was fixed for an acceptable Word Error Rate (WER). In both cases, however, the WER significantly exceeded the threshold. The interview with Lidia produced a 37.9% WER, while the one involving Lina and Nella reached a WER of 43% (WER was obtained using the *wersim* package in R). Nevertheless, the adoption of the ASR during the transcription seems to provide valuable help in simplifying the process. The technical limits reported here may be overcome in the (near) future, as well as the economic ones.

On the other hand, legacy data pose huge issues with respect to the quality of the signal because: *a*) in most cases, only a single microphone, placed at a considerable distance from the main speaker, is used, resulting in several speakers in the same audio track, i.e., no channel separation; *b*) practical constraints play a role with regard to respecting audio quality standards during the recording of the interview (unexpected factors); and *c*) the interview style (overlappings, changes in volume, vernacular forms).

In the light of these issues, some changes in the workflow are currently being carried out: i.e., the addition of further chain elements intended for audio definition and noise reduction in order to enhance the performances of the T-Chain. At the same time, the potential of ASR with such data should be emphasized. In case we succeed in collecting a considerable number of interviews from Ravensbrück, we could be able to compute a new language model which will be undoubtedly more effective for this type of legacy data. At present, however, both a manual correction phase for the ASR results and a full-manual transcription for the most complex audio segments remain the only feasible procedure to get the optimal transcription.

In view of the objective to facilitate the use and re-use of the four interviews for various disciplines, the verbatim transcriptions of the digitised audio have been created in such a way as to adhere to the requirements of linguistic research. To this end, the actual dialogues' diarization has been transcribed as it is, even in case of non-linear exchange (e.g., speech overlaps or abnormal turn-takings). Hesitations and reformulations have been transcribed as well, together with other salient acoustic signals (e.g., interjections, laughing). In the transcript, both the discontinuities produced by the actors (e.g., vocalizations like *eeh*, *mmm*, and silences) and the ones produced by contextual factors (e.g., environmental noises, technical issues of the record) have been included. Following such criteria, approximately 1025 minutes have been manually transcribed. The software used to create most of the manual transcript was [OCTRA](#)[4] (ver. 1.4.3). In some cases, the open-source Audacity software (ver. 3.0.5) was used to increase and decrease the signal amplitude to solve particularly complex speech segments. The transcription was done on an ASUS ZenBook Pro15 laptop running Windows 10.

## 3    Ravensbrück Multilingual Survey

The second part of the project aimed at reaching out to existing oral history archives and to authors who have used interviews for their publications on Ravensbrück in languages other than Italian, to explore the possibilities for contributing to a multilingual resource family (Calamai et al., 2021). This means that we have also identified some 'uncurated' material that stems, in some cases, from the pre-digital era, such as was the case with the Bruzzone archive. Although the digitisation of this material is beyond the scope of this project and the mission of CLARIN, it is important for the various research communities to know that such material exists. Consequently, in the near future, we will connect to the staff of [EHRI](#)[5] (European Holocaust Research Infrastructure) to inform them about our findings and endeavours to trace oral history data on Ravensbrück.

In the Netherlands, seven authors of books on Ravensbrück have been traced and contacted through the website of the [Ravensbrück committee](#)[6], to find material that is not yet published online. In addition,

---

[4] https://clarin.phonetik.uni-muenchen.de/apps/octra/octra/login
[5] https://www.ehri-project.eu/
[6] https://www.ravensbruck.nl/

the museum of resistance in Amsterdam and the broadcasting company VPRO have also been contacted. With regard to existing online oral history collections, the project can draw on [Getuigenverhalen.nl](#)[7], which is already directly accessible online. Concerning interviews conducted in English, there is a vast array of online interviews projects from which data could be harvested for the resource family: Shoah Visual History (US), Fortunoff Collection (US), United States Holocaust Memorial Museum (US), Imperial War Museum (GB). The USHMM is the only institute to provide direct online access to metadata, audio/video and transcripts, and has expressed its interest in collaborating and sharing its resources. In other cases, different forms of controlled access have been encountered. With regard to interviews held in German, we have found three online archives to draw on: one [German large-scale video-multilingual archive](#)[8] initiated by film-maker Loretta Walz, who is interested in collaboration, and two Austrian oral history collections, Erzählte Geschichte and VideoArchiv-Projekt Ravensbrück. To explore the possibility to broaden the range of languages, we have also identified archives in [Poland](#)[9] and [Spain](#)[10].

With regard to the diversity of variables that determine the ease with which the resources can be found and accessed and the level of richness of the data, the following categories can be distinguished:

a. Analogue or digitised interview data that is not available online but in private hands or at foundations run by volunteers, or held at archives, libraries, and museums, but without direct access because of lack of metadata description (e.g., Bruzzone archive before the CLARIN funded project);
b. Digitized or digital-born interview data on Ravensbrück that is part of a broader project, that abides to a metadata standard and can be easily identified through a refined search environment and can be directly accessed online (USHMM, Getuigenverhalen);
c. Same interview data of the point (b), but with access restricted to registered users after the creation of a personal account (Shoah Visual History, Fortunoff);
d. Digitized or digital-born interview that has been generated for a specific project on Ravensbrück, and are either published online as an autonomous resource (the video archive of Loretta Walz in Germany), or after some time have been integrated into a broader library system. The VideoArchiv-Projekt Ravensbrück can be found in the Austrian mediatheque, but the possibilities for granular searches with regard to the metadata of a specific interview are therefore very limited. The same applies to other projects that have been first created as autonomous entities, but as funding ends, have been integrated into a library system.

Within each category, several variables should be taken into account:

1. The modality of the interview (audio and/or video);
2. The mono- or multi-linguality of the interview;
3. The style (in-depth interviews generated according to rigorous academic standards or more casual interviews filmed at an occasion (international gathering, manifestation, current affairs program);
4. The different categories of survivors involved: political prisoners, resistance fighters, aid workers, Jews, Jehovah's witnesses, gypsies, and groups defined as 'socially deviant' (homosexuals, prostitutes, petty criminals);
5. The relationship between the interviewer and the interviewee.

The richness of perspectives offered by a multilingual resource family about Ravensbrück will offer novel points of view on language diversity within the context of digitisation and public history, conversational styles, and interview styles.

## 4 Conclusions

This paper shows how the integration of legacy data from the pre-digital era in the CLARIN infrastructure poses considerable challenges. It has been necessary to work on the creation of an *ad-hoc* CMDI profile explicitly devoted to oral history materials originally stored in analogue carrier. At the same time, legacy data can help automatic speech transcription (especially when they appear to be

---

[7] [https://www.getuigenverhalen.nl/](https://www.getuigenverhalen.nl/)
[8] [https://videoarchiv-ravensbrueck.de/de](https://videoarchiv-ravensbrueck.de/de)
[9] [https://audiohistoria.pl/](https://audiohistoria.pl/)
[10] [https://ajuntament.barcelona.cat/arxiumunicipal/arxiuhistoric/en](https://ajuntament.barcelona.cat/arxiumunicipal/arxiuhistoric/en)

thematically coherent and in a good state of preservation). The project *Voices from Ravensbrück* produced full verbatim transcription of the Italian interviews collected by Anna Maria Bruzzone: a valuable resource for a number of social sciences and humanities sciences, especially in view of its upcoming storage in an accessible CLARIN repository. In parallel, potential partners and suitable data for creating a multilingual collection of existing oral history interviews with survivors of concentration camp Ravensbrück were identified and contacted in order to create a resource family that is compliant with CLARIN standards, enabling scholars to analyse interviews from a comparative multilingual and multidisciplinary perspective. This last aspect, in particular, directly recalls the transnational perspective of the project: at the moment, scholars from five distinct European institutions (specifically from Italy, The Netherlands, and Germany) are involved in the project, each one bringing specific expertise and competencies.

## Acknowledgements

## References

Beccaria Rolfi, L. and Bruzzone, A. M. 2016. *Als Italienerin in Ravensbrück. Politische Gefangene berichten über ihre Deportation und ihre Haft im Frauen-Konzentrationslager*. Metropol Verlag, Berlin.

Beccaria Rolfi, L. and Bruzzone, A. M. 2020. *Le donne di Ravensbrück. Testimonianze di deportate politiche italiane*. Einaudi, Turin.

Calamai, S., Beeken, J., Henk Van Den Heuvel H., Broekhuizen, M., van Hessen A., Draxler C. and Scagliola S. 2021. "Voices from Ravensbrück. Towards the Creation of an Oral and Multilingual Resource Family". In Monachini, M. and Eskevichm M. (Eds.), *Proceedings of CLARIN Annual Conference 2021*. Virtual Edition: 16-19. https://office.clarin.eu/v/CE-2021-1923-CLARIN2021ConferenceProceedings.pdf.

Calamai, S., Kolletzek C. and Kelli, A. 2019. "Towards a protocol for the curation and dissemination of vulnerable people archives". In Skadina, I. and Eskevich, M. (Eds.), *Selected papers from the CLARIN Annual Conference 2018*. Linköping University Electronic Press, Linköpings: 28-38. https://ep.liu.se/ecp/159/003/ecp18159003.pdf.

Draxler, C., Van den Heuvel, H., Van Hessen, A., Calamai, S., Corti, L. and Scagliola, S. 2020. "A CLARIN Transcription Portal for Interview Data". In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC2020)*. Virtual Edition: 3346-3352. http://www.lrec-conf.org/proceedings/lrec2020/pdf/2020.lrec-1.411.pdf.

Frisch, M. 1990. *A Shared Authority: Essays on the Craft and Meaning of Oral and Public History*. New York University Press, New York.

Hogervorst, S. 2010. *Onwrikbare Herinnering. Herinneringsculturen van Ravensbrück in Europa*. Uitgeverij Verloren, Hilversum.

Kisler, T., Schiel, F. and Sloetjes, H. 2012. "Signal processing via web services: the use case WebMAUS". In *Proceeings of Digital Humanities Conference*, Virtual Edition: 30-34. https://pure.mpg.de/rest/items/item_1850150/component/file_1850149/content.

Scagliola, S., Corti, L., Calamai, S., Karrouche, N., Beeken, J., Van Hessen, A., Draxler, Chr., Van den Heuvel, H., Broekhuizen, M. and Truong, K. 2020. "Cross disciplinary overtures with interview data: Integrating digital practices and tools in the scholarly workflow". In Simov, K. and Eskevich, M. (Eds.), *Selected Papers from the CLARIN Annual Conference 2019*. Linköping University Electronic Press, Linköpings: 126-136. https://doi.org/10.3384/ecp2020172015.

# Appendix 1

## Specimen of metadata profile (Interview with Lidia Beccaria Rolfi)

| Component | Component I | Component II | Element | Value |
|---|---|---|---|---|
| **ResourceFamily** | | | | |
| | | | Description | Anna Maria Bruzzone Archive - Ravensbrück series |
| | | | ID | *to be determined* |
| **InterviewGeneral** | | | | |
| | | | NumberOfSpeaker | 2 |
| | | | Duration | 06:56:58 |
| | | | Owner | Università degli Studi di Siena, Archivio Storico dell'Ospedale Neuropsichiatrico di Arezzo, Arezzo (AR), Italy |
| | | | Genre | Interview |
| | | | Modality | Audio |
| | **ContextOfCreation** | | | |
| | | | ContextOfCreation | This interview with the Ravensbrück ex-deported Lidia Beccaria Rolfi was originally recorded on a series of analog cassette tapes by the oral historian Anna Maria Bruzzone in preparation of her book (Women of Ravensbruck; original title: Le Donne di Ravensbruck, Einaudi, first publication: 1978). The interview comes from the union of 9 different files, deriving from the digitisation of as many sides of 5 audiocassettes. The original audiocassettes are located in the Archivio Storico dell'Ospedale Neuropsichiatrico di Arezzo (Università degli Studi di Siena), Arezzo (AR), Italy. |
| | | **Contact** | | |
| | | | Role | Coordinator of the project |
| | | | Name | Silvia |
| | | | Surname | Calamai |
| | | | Organisation | Università degli Studi di Siena, Siena, (SI), Italy |
| | | | E-mail | silvia.calamai@unisi.it |
| | | **Contact** | | |
| | | | Role | Curator of the analogue collection |
| | | | Name | Lucilla |
| | | | Surname | Gigli |
| | | | Organisation | Università degli Studi di Siena, Siena, (SI), Italy |
| | | | E-mail | lucilla.gigli@unisi.it |

| Component | Component I | Component II | Element | Value |
|---|---|---|---|---|
| | **ContextOfDigitisation** | | | |
| | | | ContextOfDigitisation | The digitisation of the audiocassettes was carried out on 10/01/2019 at the Centro di Sonologia Computazionale (Università di Padova), Padova (PD), Italy. |
| | | **Contact** | | |
| | | | Role | Expert in charge |
| | | | Name | Alessandro |
| | | | Surname | Russo |
| | | | Organisation | Centro di Sonologia Computazionale, Dipartimento di Ingegneria Informatica (Università di Padova), Padova (PD), Italy |
| | | | E-mail | alessandro.russo@unipd.it |
| | **DigitalAccess** | | | |
| | | | DigitalAccess | *to be determined* |
| | | | Availability | *to be determined* |
| | | | CatalogueLink | *to be determined* |
| | **InterviewSummary** | | | |
| | | | InterviewSummary | [0.00 – 15.00] Interview starts in the middle. Lidia says that when she was born, her father was 45 years old. She had a very |
| | | | | [0.00 – 15.00] Lidia resumes with her memory of July 25 as an illusion that the war was over, with one brother in military |
| | | | | [0.00 – 15.00] Lidia talks about her experience in jail. The first night she was with Carletti and Pina Doleatti, but since they |
| | | | | [0.00 – 15.00] Lidia says that in the blocks, there were no toilets. After an interruption, the tapes resumes: Lidia recounts |
| | | | | [0.00 – 15.00] The taping resumes from the episode of the Christmas meal. They were punished and then transferred to |
| | | | | [0.00 – 15.00] She continues talking about the day of the evacuation. They are told to take everything they need. She takes |
| | | | | [0.00 – 15.00] Lidia continues to talk about when they wanted her to get off the train even though she had the same pass |
| | | | | [0.00 – 15.00] The following year, she returned to France to meet her companions because she knew the addresses by |
| | **Language** | | | |
| | | | Language | Italian |
| | | | Iso-369-3-code | ita |
| | | **Multilinguality** | | |
| | | | Multilinguality | Monolingual |
| | **Interviewee** | | | |
| | | | Name | Lidia |
| | | | Surname | Beccaria Rolfi |
| | | | BirthPlace | Mondovì (CN), Italy |
| | | | BirthCountry | Italy |
| | | | ResidentPlace | Mondovì (CN), Italy |
| | | | ResidentCountry | Italy |
| | | | Role | Former deportee in the Ravensbrück concentration camp, writer, anti-fascist activist |
| | | | Family | Parents employed as farmers. Last of six siblings (of which known Rita, Luigi, and Enrico). Mother of one (Aldo Rolfi). |
| | | | EthnicGroup | Italian (Piedmontese) |
| | | | Age | 51 (at the time of the interview). Deceased in 1996. |
| | | | BirthYear | 1925 |
| | | | Sex | Female |
| | | | Education | Teaching diploma |
| | | | Profession | Teacher, writer |
| | | | Anonymized | False |
| | | **Language** | | |
| | | | LanguageName | Italian |
| | | | Iso-369-3-code | ita |
| | | | Description | The spoken variety is the Regional Piedmontese Italian. Local lexical forms are sometimes adopted. Other languages |

| Component | Component I | Component II | Element | Value |
|---|---|---|---|---|
| | Interviewer | | | |
| | | | Name | Anna Maria |
| | | | Surname | Bruzzone |
| | | | Role | Main investigator, historian |
| | | | RelationToInterviewee | Shared native city (Mondovì) |
| | | | RelationToProject | Author, main investigator |
| | | | BirthPlace | Mondovì (CN), Italy |
| | | | BirthCountry | Italy |
| | | | ResidentPlace | Torino (TO), Italy |
| | | | ResidentCountry | Italy |
| | | | Family | unknown |
| | | | EthnicGroup | Italian |
| | | | Age | 51 (at the time of the interview). Deceased in 2015. |
| | | | BirthYear | 1925 |
| | | | Sex | Female |
| | | | Education | Bachelor of Arts, Major in Psychology |
| | | | Profession | Teacher, historian, writer |
| | | | Anonymized | False |
| | | Language | | |
| | | | LanguageName | Italian |
| | | | Iso-369-3-code | ita |
| | | | Description | The spoken variety is the Regional Piedmontese Italian. |
| InterviewContent | | | | |
| | InterviewKeywords | | | |
| | | | Keywords | Mondovì; First postwar period; Alpini corps; Fascism; Mussolini; racial laws; Spanish Civil War; Abyssinian War; Second |
| | Full transcript | | | |
| | | | FullTranscript | Yes |
| | Coverage | | | |
| | | | SpatialCoverage | Northern Italy, Ravensbrück |
| | | | TimeCoverage | 1918-1945 |
| InterviewMethod | | | | |
| | | | RecruitmentMethod | unknown |
| | | | PreInterviewInformatio | unknown |
| | | | TypeOfInterview | Free interview |
| | | | TopicList | unknown |
| InterviewAudio | | | | |
| | | | AudioFileName | BRZTO061a.wav, BRZTO061b.wav, BRZTO062a.wav, BRZTO062b.wav, BRZTO063a.wav, BRZTO063b.wav, BRZTO064a.wav, |
| | | | AudioFormat | .wav |
| | | | AudioQuality | Good |
| | | | RecordingConditions | Indoors |
| SpeechTecnicalMetad | | | | |
| | | | SamplingFrequency | 96 kHz |
| | | | NumberOfChannel | 2 |
| | | | ByteOrder | little_endian |
| | | | Compression | none |
| | | | BitResolution | 24 bit |
| | | MimeType | | |
| | | | MimeType | audio/wav |
| InterviewAnnotation | | | | |
| | | | AnnotationProtocol | time aligned transcript |
| | | | CharacterEncoding | ASCII |
| | | | AnnotationFileName | BRZTO061a.txt, BRZTO061b.txt, BRZTO062a.txt, BRZTO062b.txt, BRZTO063a.txt, BRZTO063b.txt, BRZTO064a.txt, |
| | | | AnnotationType | ortographic, phonetic |
| | | | Standards | [Lidia:] speaker |
| | | | AnnotationFormat | .txt |