THE GRA.FO PROJECT AND THE NEW GRA.FO PANEL: TOOLS FOR DIGITIZATION, STORAGE AND CATALOGUING OF ORAL ARCHIVES

Irene Ricci[1], Chiara Bertini[1], Francesca Biliotti[2], Silvia Calamai[2], Dei Rossi[3]

[1] Scuola Normale Superiore, Pisa
[2] Università degli Studi di Siena, Arezzo
[3] WebSoup s.n.c., Lucca

The cultural heritage preserved in some oral archives risks being loss because of the deterioration of the carriers and the lack of methodological tools for conservation (Canazza et al. 2012). *Grammofoni. Le soffitte della voce* (*Gra.fo*), a project conducted by Scuola Normale Superiore of Pisa & University of Siena and financed by Regione Toscana (PAR FAS 2007-2013), spotted, collected, digitised, catalogued, and transcribed oral material recorded by scholars and amateurs in Tuscany. *Gra.fo* provides an incredibly rich repository of Tuscan oral documents that have so far remained invisible: tales, proverbs, songs, interviews, ethno-texts, linguistic questionnaires, and much more (Calamai et al. 2013).

So far, more than 2800 hours of recording occupying nearly 2000 carriers (reel tapes and compact cassettes) have been digitized. Yet, thanks to funding and active help from organizations and private individuals, the collection of catalogued oral documents accessible via the *Gra.fo* web portal (http://grafo.sns.it/) is ever-growing.

The sound documents, edited from the content of the carriers (on the basis of the information contained in the recording, or in the accompanying material, or written on the carrier) and called 'units of audio consultation' (see *infra*), are put in relation with the original documents from which they derive thanks to a database developed in MySQL which is populated by means of two *ad hoc* applications, *Grafo Preservation Panel* and *Grafo Panel*, which make up and integrated system for the management of the *Gra.fo* digital archive. These tools allow for the homogenization of the information and data related to the preservation copies[1] stemming from different archives, in compliance with the European standards and through a systematic, coherent, detailed workflow.

The work aims at presenting the upgrade undergone by the two above-mentioned informatics tools.

Grafo Preservation Panel

Grafo Preservation Panel is a stand-alone application developed with Java that computerizes the preservation process by allowing a series of semi-automated procedures for generating a preservation copy, describing it by entering all the necessary information (resulting from the digitization process and emerging from the metadata) in the database, managing the archive via the distribution of the information and with the help of quality-check tools (as is the case with the collocation of all documents within each preservation copy), and so on.

In addition, *Grafo Preservation Panel* carries out the transfer of the preservation copies to a server for long-term conservation.

Grafo Panel

Grafo Panel (http://grafopanel.sns.it/) is a web application that enables the cataloguing of oral documents following a protocol established by the *Gra.fo* research group (Calamai, 2012; Calamai, Bertinetto 2014). The cataloguers (registered and authorized for the purpose) use this application to create and fill in the cataloguing records related to the units of audio consultation, which correspond to single linguistic events contained in the relative preservation copies[2]. The application allows

---

[1] An organised set of data containing all the information present in the carrier together with its description and the documentation of the conservation process.

[2] While the preservation copy faithfully reproduces the original recording without considering its content, the 'unit of audio consultation' derives from an interpretive process in which the cataloguer edits the original file and creates as many different documents as the recorded events are.

cataloguers to enter information such as genre and topic, to edit a summary of the document, to associate it with the relevant keywords and with the relevant interviewers and interviewees among those catalogued in the software, to upload the related .pdf files (transcriptions and accompanying materials), and much more.

In describing the units, cataloguers are helped by the presence of pull-down menus and text fields that allow for uniformity of labelling and limit human errors. In addition, *Grafo Panel* has some automated quality-check systems: for example, the creation of a new cataloguing record is impossible without entering some specific information (title, language variety, archive to which the document belongs). The upgrade of the software has brought about some improvements in the interface, which is now more user-friendly, and, more importantly, in the performance of certain operations. For example, while in the previous version the editing of the existing keywords and of the existing information regarding interviewers and interviewees was restricted to the administrator of database, now the operation can be carried out by the cataloguers and is thus less complicated and less time-consuming. Another important improvement is the introduction of restrictions to the names of the files that can be associated to the cataloguing records: since each cataloguing record has a specific code and all the related files (.mp3 file containing the unit of audio consultation and .pdf files containing transcriptions and accompanying materials) must be named with the same code, the software only accepts files bearing the correct name. In addition, the new software does not allow a cataloguer to log in as a fellow-cataloguer, a possibility given by the previous version: now cataloguers can access the cataloguing records of their colleagues, but cannot modify them, which makes the whole procedure safer.

All the information recorded in *Grafo Panel* converge in the same database containing the information recorded in *Grafo Preservation Panel* about the original carriers. This ensures a safe connection between the data derived from the digitisation process and those produced within the cataloguing process.

Grafo Preservation Panel and *Grafo Panel* are two informatics tools optimised for the automation of the processes of preservation (creation of preservation copies) and cataloguing (description of single oral documents). They present two main advantages: 1. they allow for the check of data consistency and 2. they are time-saving with respect to systems based on manual compilation.

References

Calamai, S. (2012), Ordinare archivi sonori: il progetto Gra.fo, Rivista Italiana di Dialettologia, 35.

Calamai S., P.M. Bertinetto (2014), Le soffitte della voce. Il progetto Grammo-foni, Vecchiarelli, ISBN 88-8247-352-5.

Calamai S., F. Biliotti, P.M. Bertinetto, C. Bertini, I. Ricci, G. Scuotri (2013), "The Gra.fo sound archive: architecture, methods and purpose", in A.C. Addison, G. Guidi, L. De Luca, S. Pescarin (eds.), Proceedings of the 2013 Digital Heritage International Congress (DigitalHeritage), Institute of Electrical and Electronics Engineers Inc., 2013: 439 [ISBN: 978-1-4799-3169-9].

Canazza S., S. Calamai, P.M. Bertinetto, A. De Dominicis (2012), "A protocol for the preservation of speech documents archives: towards the digital curation of the Carta dei Dialetti Italiani", in Proceedings of the 5th International Congress Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin, Istanbul, Turkey, 22nd-25th November 2011, vol. IV, pp. 84-93.