

Bulletin de l'AFAS

Sonorités

45 | 2019 :
Bulletin de l'AFAS. Sonorités n° 45
Présentation de corpus

Web dissemination of Tuscan oral archives

Diffusion sur le Web des archives orales de la Toscane

FRANCESCA BILIOTTI ET SILVIA CALAMAI

Résumés

Diffusion sur le Web des archives orales de la Toscane

Le projet Grammo-foni. *Le soffitte della voce (Gra.fo)* a été réalisé par l'École Normale Supérieure de Pise et l'Université de Sienne avec le soutien financier de la Région Toscane. Il valorise une trentaine de collections sonores recueillies sur l'ensemble de la région par des chercheurs et des passionnés de la tradition orale. La valorisation de ces archives – jusqu'ici inconnues du grand public – s'effectue sur plusieurs étapes qui vont de la découverte de ces archives à leur numérisation (et leur restauration, si nécessaire) en passant par le catalogage des enregistrements de chaque collection. Les documents sonores sont accessibles par un portail web (<http://grafo.sns.it>), qui permet aux témoins enregistrés et aux enquêteurs d'accéder aux entretiens enregistrés ainsi qu'à leur analyse, leurs transcriptions et tous les documents qui ont accompagné l'enquête le cas échéant. Cet article décrit le portail web de *Gra.fo* et ses diverses possibilités de recherche jusqu'aux notices dans le catalogage.

The project *Grammo-foni. Le soffitte della voce (Gra.fo)*, carried out by Scuola Normale Superiore and the University of Siena with funding from Regione Toscana (PAR FAS 2007-13), preserved around thirty oral archives collected within various fields of research by scholars and amateurs in the

Tuscan territory. The preservation of such archives, which have so far remained unknown to the large public, entailed their detection as a first step, and then the digitisation (including restoration, when necessary) and cataloguing of the recordings contained in them. The oral documents preserved are disseminated via a web portal (<http://grafo.sns.it>) that allows registered users to access the audio files and the corresponding cataloguing records, together with the relative transcriptions and accompanying material (when available). The present paper describes the *Grafo* web portal: from the types of search provided to the cataloguing records.

Entrées d'index

Mots-clés : phonothèque, conservation, traitement documentaire, archives en ligne, Toscane, Italie

Texte intégral

Introduction

- 1 The *Grafo* web portal (grafo.sns.it) hosts an incredibly rich repository of Tuscan oral testimonies that have so far remained invisible and inaudible: tales, proverbs, songs, interviews, ethnотexts, linguistic questionnaires, biographies, and much more. The portal offers the possibility to access 2208 documents, corresponding to 2800 hours of recording, containing the voices of little less than 300 interviewees and 143 interviewers. Such stunning numbers are the result of a joint effort between Scuola Normale Superiore (SNS) and the University of Siena (UNISI). With funding from Regione Toscana (PAR FAS 2007-13), the two universities carried out a two-year project for the preservation of Tuscan oral archives that constitutes a *unicum* in the Italian panorama: the project *Grammo-foni. Le soffitte della voce (Grafo)*. *Grafo* detected, digitised (restored, when necessary), and catalogued oral documents stemming from around thirty oral archives collected within various fields of research by scholars and amateurs in the Tuscan territory.

The history of the project

- 2 *Grafo* began in 2011 after long collaboration between Pier Marco Bertinetto (SNS) and Silvia Calamai (UNISI). Crucially, the project came after an important census was carried out at regional level by a research group made of anthropologists: *I custodi delle voci*², which detected a considerable number of Tuscan oral archives. But

Grafo also aimed at going beyond disciplinary boundaries and include linguistic archives, which were hardly considered in previous censuses. For this purpose, an interdisciplinary team was put together: two linguists, Pier Marco Bertinetto and Silvia Calamai, were at the head of the SNS Research Unit and of the UNISI Research Unit respectively, assisted by: i) a computer scientist (Federica Bressan) who developed a software for archiving and cataloguing oral documents and dealt with the digitisation process; ii) a technician who succeeded Federica Bressan in the digitisation of the archives (Gianfranco Scuotri), iii) five cataloguers with expertise in the fields of linguistics, dialectology and anthropology who carried out the description and transcription of the oral documents preserved (Cristina Bertoncin, Francesca Biliotti, Nadia Nocchi, Luca Pesini, Valentina Zingari), iv) two physics as technical and administrative staff (Chiara Bertini and Irene Ricci). In addition, SNS allocated a specific contract to an external company for the creation of the web portal. *Grafo* officially ended in 2013, but still survives thanks to the project *Voci da ascoltare*, conducted by the University of Siena with partial funding from Unicoop Firenze, which is taking the work a step further by designing educational projects and activities based on the digital oral documents stemming from *Grafo*.

The collections

- 3 *Grafo* preserved different kinds of archives stemming from different fields of research. Some of the archives preserved originated from research projects conducted by linguists in order to document or investigate specific features of Tuscan dialects (e.g. 'Alto Mugello', 'Atlante Lessicale Toscano', 'Silvia Calamai', 'Carta dei Dialetti Italiani', 'Seminari di Linguistica Generale'). Other archives were collected in an anthropological, folkloric or ethnomusicological perspective and thus concern folk music, folk literature and folk culture in general (e.g. 'Edda Ardimanni', 'Roberta Beccari', 'Vanna Brunetti', 'Anna Buonomini', 'Paolo De Simonis' - collection 'Canti popolari del Mugello', 'FLOG' - collections 'Gilberto Giuntini' and 'Nunzi Gioseffi', 'Sergio Gargini', 'Benozzo Gianetti', 'Duse Lemetti - Gruppo Vegliatori', 'Museo del Bosco'). Yet others stem from history and sociology, and concern topics like working conditions in the twentieth century, labour movement, women labour, "Italian diaspora", impact of industrialisation on rural society, memories of the First and Second World War (e.g. 'FLOG' - collection 'Andrea Grifoni', 'ASMOS', 'Neri Binazzi', 'Cappelli di paglia', 'Dina Dini', 'Elba', 'Roberto Segnini', 'Angela Spinelli').

How to search documents on the *Gra.fo* web portal

- 4 Two levels of accessibility are envisaged: one for registered users and one for unregistered users³. The latter can access the descriptions of the archives and the cataloguing records of the single documents, but cannot access nor download the relative audio files, transcriptions, and accompanying materials. In order to register to the website, a user must click on '*Registrali*' ('Register'), provide some personal data (name, surname, date of birth, sex, username, e-mail address) and a captcha code, click on '*Salva*' ('Save'), and wait for a confirmation e-mail from grafo@sns.it. Once the registration is completed, the user can have complete access to the resources available in the portal.
- 5 At the moment, the *Gra.fo* web portal is hosted by the SNS website and thus follows SNS privacy policies, but a parallel migration to the CLARIN-it repository is being pondered (<http://clarin-it.it>), together with the possibility of searching data from the *Gra.fo* archive via the Virtual Language Repository (VLO)⁴.
- 6 In the *Gra.fo* web portal, the documents are accessible via two different types of search: Linguistic-area search and Advanced search. As shown in Fig. 1, in the Linguistic-area search, the user is presented with an interactive linguistic map of Tuscany based on the taxonomy proposed by Giannelli⁵. By clicking on any area of the map, the user will get a record enlisting the archives and the number of single documents referring to that particular area (e.g. in Fig. 1, "Brunetti Vanna 491" means that 491 documents belonging to *Archivio "Brunetti Vanna"* contain speech of the Florentine area). At that point, it is possible to access the corresponding cataloguing records by simply clicking on one of the results included in the list, or to refine the research by clicking on '*Raffina ricerca*' ('Refine the search'), a link that takes directly to the Advanced search page.

Fig. 1. Linguistic-area search⁶



- 7 As shown in Fig. 2, in the Advanced search, it is possible to search documents according to three main categories: 'Contenti' ('Content'), 'Luoghi' ('Place'), and 'Tempi' ('Time'), which in turns include some subdivisions. The first category, 'Contenti', includes 'Argomento' ('Topic'), 'Genere' ('Genre'), 'Tipologia' ('Type'), and 'Varietà linguistica' ('Language variety'):

- 'Argomento' – Here the user can choose among around 130 different topics (among others, Agriculture, Anarchism, Animals, Art, Autobiographies, Biographies, Blacksmiths, Carnival, Cinema, Clothing, Coalmen, Cutlers, Dialects and language varieties, Domestic activities, Drug addiction, Emigration, Environment, Exhibitions, Family, Fascism, Fishing, Folk dance, Folk literature, Folk medicine, Folk music and songs, Folk theatre, Folk traditions, Food, Games, Handicraft, Human body, Immigration, Legends, Literature, Local history, Magic, Material culture, Music festivals, Nazism, Peasant culture, Peasant traditions, Political history, Politics, Post-war period, Pre-industrial society, Prostitution, Racism, Religion, Religious feasts, Rituals, School, Sharecropping, Theatre, Time, Traditional family, Traditional festivals, Traditional food, Traditional jobs, Traditions, Women's condition, Women's history, Work, 1st World War, 2nd World War). It is possible to search different topics simultaneously: since only one topic per document is assigned, the results will include all the documents containing one of the selected topics.
- 'Genere' – Here the user can choose among around forty different genres (among others, Answer to linguistic questionnaire, Autobiography, Ethno-text, Image/object description, Interview, Legend, Lullaby, Narrative song,

Poem, Political song, Prayer, Proverb, Reading, Recipe, Religious poetry, Riddle, Ritual, Spontaneous conversation, Tale, Theatre, Tongue twister). Like with the topics, it is possible to search different genres simultaneously: since only one genre per document is assigned, the results will include all the documents appointed with one of the selected genres.

- *'Tipologia'* – This section includes two different fields, the second being dependent on the first. In the first, the user can choose between sung, non-sung, and mixed documents (the latter containing both speech and singing). Sung documents can in turn be distinguished in formalised documents (e.g. Lullaby, Narrative song), improvised documents (e.g. improvised *ottava rima*), and mixed documents (characterised by both formalised and improvised singing). Similarly, non-sung documents can in turn be distinguished in formalised documents (e.g. Poetry, Proverb), non-formalised documents (e.g. Interview, Spontaneous conversation), and mixed documents (characterised by both formalised and non-formalised speech). When choosing mixed documents in the first field, there is no further option available in the second (i.e. mixed documents cannot be distinguished into further categorisations).
- *'Varietà linguistica'* – This field allows the user to search documents according to the language variety spoken by the informants. It is possible to choose different language varieties simultaneously: since only one variety per document is assigned, the results will include all the documents containing one of the selected varieties. The taxonomy adopted by *Grafo* stems from Luciano Giannelli's taxonomy⁷ and includes urban varieties (of Florence, Prato, Pistoia, Lucca, Massa, Pisa, Leghorn, Arezzo, Siena, Grosseto), areas of influence (of Florence, Pistoia, Lucca, Pisa, Leghorn, Arezzo, Siena, Grosseto), areas of transition (of Volterra, Massa, Piombino), and other minor varieties (e.g. of the Elba Island). The sociolinguistic motivations for such choice are twofold: a) cities are a vehicle for linguistic identity and usually influence the surrounding areas; b) Tuscany does not have a hegemonic centre influencing the whole territory of the region.

⁸ *'Luoghi'* refers to i) the city where the archive comes from and ii) the city where the single documents were collected, as the two might not correspond (e.g. *Archivio "Carta dei Dialetti Italiani"* comes from Florence, but its documents were collected in every Tuscan municipality⁸). In both fields, the user can see the full list of Tuscan cities and select one or more of them.

- 9 *'Tempi'* refers to the time when the single documents were collected. By setting the dates on the available calendars, the user can search for documents collected over a certain period of time (e.g. between 1 January 1970 to 31 December 1990).
- 10 In the Advanced search, the user can select only one criterion (e.g. search all sung documents, or documents collected in Florence, or documents about the Second World War), or use different criteria simultaneously (e.g. one can search for sung, formalised documents produced in the Florentine variety, or for interviews about labour work collected in Prato between 1970 and 2000).
- 11 After selecting the desired criteria, the user will get a record enlisting the archives and the number of single documents satisfying those particular criteria (e.g. the list on the right in Fig. 2 shows the names of the archives containing interviews collected from 1961 to 2000, together with the number of relevant documents). By clicking on any of such items, the user can access the related cataloguing records.

Fig. 2. Advanced search

CONTENUTI		Archivio	Unità
Argomento: <input type="text"/>		Becchi Roberta	16
Genere: <input type="text" value="intervista"/>		Stassi Neri	20
Tipologia: <input type="text" value="Selezione cantato"/> <input type="text" value="Selezione primo cantato"/>		Calamai Silvia	57
Area Linguistica: <input type="text" value="Selezione zona"/>		Spinelli Angela	83
LUGHI		Dini Dini	7
Pratica presentata a nome: <input type="text" value="Selezione provincia"/> Pratica acquisite in nome: <input type="text" value="Selezione provincia"/>		Seminari di Linguistica Generale	0
TEMPI		Cianetti Benozzo	7
Anno di riferimento: Dal <input type="text" value="01-01-1961"/> * <input type="text" value="31-12-2000"/>		Lemetti Duse - Gruppo Vegetariani	10
		Andimanni Edda	8
		FLOG	1
		Museo del Bosco	19
		Gargini Sergio	18
		Buonomini Anna	18
		ASMOS	9
		Alto Mugello	17
		Cappelli di paglia	33
		Elba	6
		Strutini Vanna	11
		De Simonia Paolo	4
		Unione Industriale Prato	0
		Storia di Prato - Alessandro Pescarolo	12
		Seghini Roberto	0
		Carte dei Dialetti Italiani	0


Cataloguing records

12 As shown in Fig. 1 and 2, the result of both the Linguistic-area search and the Advanced search is a list of items presenting the names of the archives and the number of relative documents satisfying the search criteria. By clicking on one of the items, the user will access a page showing the relevant documents positioned along a timeline according to their date of collection. At this stage, by clicking on one of the titles positioned along the timeline, the user can see a preview of the related cataloguing record (Fig. 3) bearing the following information: title, summary, date of collection, name of the archive (and relative subsections⁹) to which it belongs, audio file (that can be played, but not downloaded from the preview). It is possible to move from one preview to the next (or preceding) one by using the left and right arrows positioned on the left and right margins of the page.

Fig. 3. Preview of a cataloguing record



The complete cataloguing record is accessible via the link 'Dettagli' ('Details') and shows all the information available for the selected document organised in two vertical sections (Fig. 4). On the left, there are the title, the full summary, and the name of the archive (and relative subsections) to which the document belongs (full description of archive and subsections appear when moving

the mouse over the  symbols). On the right, the user can find the following: subtitle, keywords, date and place of collection, setting, type of document, topic, genre, language variety, bibliography, aims of the investigation, type of carrier, information about privacy restrictions, about interviewer(s) and interviewee(s), and about accompanying materials. On the top right, there are the downloadable files associated with the document: the audio file



- 13 Public involvement is highly welcomed in *Grafo*, which attempted to employ crowdsourcing by soliciting contributions from the web portal users. As shown in Fig. 5, these are invited 1) to entrust their oral documents to the project in order to digitise them and ensure their long-term preservation, and 2) to catalogue or transcribe one (or more) of the *Grafo* documents that have not been catalogued and transcribed yet. Those who do this will see their contribution acknowledged in the digital files and records related to the documents that they donated, catalogued or transcribed. Hence, not only does *Grafo* return to the community a treasure of immense value; it also encourages the community to take part in building it.
- 14 Unfortunately, though, the *Grafo* web portal does not seem to be able to work as a crowdsourcing platform. This is probably because the initiative has not been publicised enough and through the right media (the project is not present in any social network, which might be a limit). More generally, the portal seems to be appreciated and used mostly by researchers, less by the wider public, which finds it not very easy to consult. Therefore, the resources have been scarcely reused: apart from the dissertations supervised by Silvia Calamai, the only documented case is a documentary exhibition on the archive 'Cappelli di paglia' put up by the Soprintendenza archivistica e bibliografica della Toscana. In addition, even when some of the *Grafo* materials are reused, those who use them do not feel it their duty to inform the *Grafo* working group. The project *Voci da ascoltare* was created precisely to remedy this situation, i.e. to share the materials with a wider public (schoolteachers and students) and to ensure their correct use.

Dissemination

- 15 *Grafo* started with a campaign to raise awareness among archives' owners on the need for preserving, archiving and describing oral documents. Those who accepted to join *Grafo* were informed about the project stages and the creation of the web portal, they were asked to choose the access policy for their materials, and were invited to the *Grafo* final conference to be consigned a certificate declaring them as 'Protectors of oral culture' (Fig. 6).

Fig. 6. Leaflet and Programme of the *Grafo* final conference



Notes

- 1 S. Calamai, F. Biliotti, L. Pesini, P.M. Bertinetto, "Building an open sound archive: the case of the Gra.fo project", in Proceedings of the 6th international congress on "Science and Technology for the safeguard of Cultural Heritage in the Mediterranean Basin", Rome, vol. 3, 2014: 264-269 [ISBN 978-88-97987-05-5]; S. Calamai, 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].
- 2 Alessandro Andreini & Pietro Clemente, *I custodi delle voci. Archivi orali in Toscana: primo censimento*, Regione Toscana, Firenze, 2007.
- 3 For a detailed account on the different types of consultation provided by *Grafo* depending on the presence of confidential data in the documents, please see S. Calamai, V. Ginouvès, P.M. Bertinetto, "Sound Archives Accessibility" in Karol Jan Borowiecki, Neil Forbes & Antonella Fresa (eds.) *Cultural heritage in a changing world*, Springer: Berlin, 2016, p. 37-54.
- 4 S. Calamai & F. Frontini, *Not quite your usual kind of resource. Gra.fo*

and the documentation of Oral Archives in CLARIN. In Proceedings of the CLARIN Annual Conference 2016. Aix-en-Provence, 2016, <https://hal.archives-ouvertes.fr/hal-01395027/document>

5 Luciano Giannelli, "Italienisch: Areallinguistik VI Toskana", in G. Holtus, M. Metzeltin, C. Schmitt (eds.), *Lexikon der Romanistischen Linguistik, IV: Italienisch, Korsisch, Sardisch. Italiano, corso, sardo*, Tübingen: Niemeyer, 1988, p. 594-606.

6 The pictures show screenshots of the original Italian website; English translations are given either in note or in the text. (For the moment, the English section of the website covers only a part of the original website. A full English version of the website is a desirable future development).

7 *Ibidem* and Luciano Giannelli, *Toscana*, Pisa: Pacini, 2000.

8 For more information on the *Archivio "Carta dei Dialetti Italiani"*, please see S. Calamai, P.M. Bertinetto, "Per il recupero della Carta dei Dialetti Italiani", in T. Telmon, G. Raimondi, L. Revelli (eds.), *Coesistenza linguistica nell'Italia pre- e postunitaria*. Atti del XLV Congresso internazionale di studi della Società di Linguistica Italiana (Aosta/Bard /Torino 26-28 settembre 2011), 2 voll., Roma: Bulzoni, 2012, p. 335-356. [ISBN: 978-88-7870-722-1]; S. Calamai, F. Biliotti, "Geolinguistic archives within the Gra.fo project: Carta dei Dialetti Italiani and Atlante Lessicale Toscano", in E. Wandls-Vogt & A. Dorn (eds.), *Dialekt | dialect 2.0. Langfassungen | Long papers*. 7. Kongress der Internationalen Gesellschaft für Dialektologie und Geolinguistik (SIDG) | 7. Congress of the International Society for Dialectology and Geolinguistics (SIDG), Wien, Praesens Verlag 2017: p. 91-99.

9 For an explanation of the partition of archives in smaller sections, please see S. Calamai, P.M. Bertinetto 2014, *Le soffitte della voce. Il progetto Grammo-foni*, Vecchiarelli, ISBN 88-8247-352-5; S. Calamai 2011 "Ordinare archivi sonori: il progetto Gra.fo", *Rivista Italiana di Dialettologia*, 35, p. 135-164.

10 'COLLABORATE WITH US..... If you own sound material of linguistic, historical, anthropologic, economic interest and want to preserve it. If you are interested in cataloguing one of the documents that have not been catalogued yet. If you are interested in transcribing one of our documents and see your contribution acknowledged. CALL US at 050 509218/007 or WRITE to grafo@sns.it'.

Pour citer cet article

Référence électronique

Francesca Biliotti et Silvia Calamai, « Web dissemination of Tuscan oral archives », *Bulletin de l'AFAS* [En ligne], 45 | 2019, mis en ligne le 26 mars 2019, consulté le 08 août 2019. URL : <http://journals.openedition.org/afas/3439> ; DOI : 10.4000/afas.3439

Auteurs

Francesca Biliotti
Université de Sienne, Italie

Silvia Calamai
Université de Sienne, Italie

Droits d'auteur

Bulletin de l'AFAS. Sonorités

This site uses cookies and collects personal data.

For further information, please read our Privacy Policy (updated on June 25, 2018).

By continuing to browse this website, you accept the use of cookies. [Close](#)