



SCUOLA
NORMALE
SUPERIORE



Calamai Silvia¹, Biliotti Francesca¹, Pesini Luca³, Bertinetto Pier Marco³

BUILDING AN OPEN SOUND ARCHIVE: THE CASE OF THE *GRA.FO* PROJECT

0. Overview of the project

Oral historians and anthropologists have often underlined the urgent need to protect oral archives collected by professional scholars and ordinary people interested in dialects and ethnology [1,2]. Only recently, however, has the issue of audio materials conservation become relevant for linguists, although many scholars have gathered large amounts of audio data over the years. Fortunately, the development of *corpus* linguistics, together with the increasing interest in acoustic phonetics and audio restoration, have strongly encouraged the launch of projects aiming at collecting audio materials.

Grammo-foni. Le soffitte della voce (Gra.fo) is a two-year project jointly conducted by Scuola Normale Superiore and the University of Siena (Regione Toscana PAR FAS 2007-13). Its purpose is to detect and preserve oral documents (e.g. oral biographies, ethno-texts, linguistic questionnaires, oral literature, etc.) collected by scholars and amateurs within the Tuscan territory. *Gra.fo* aims thus at providing a first-hand documentation of Tuscan varieties from the early 1960s to the present time.

The project includes five stages:

- fostering the level of awareness on the importance of preserving this valuable (but largely invisible) product of cultural heritage;
- contacting the oral recordings' owners to legally agree for the temporary borrowing of the materials;
- collecting, digitising, and (when necessary) restoring the audio materials;
- cataloguing and partially transcribing the oral documents;
- offering the opportunity for online accessibility of digitised content for a large audience (a downloadable online database: <http://grafo.sns.it/>).

This growing repository thus will provide the opportunity to discover oral texts which, until now, have been known to a very limited number of possible users, thus ensuring the safeguarding of a specific type of endangered intangible cultural heritage. Besides, the *Gra.fo* archive offers a vast quantity of (mostly unpublished) documents for further linguistic, economic, social, political, historical, and cultural analysis.

This paper is organised as follows: § 1 presents the main oral archives collected, digitised and partially catalogued during the project; § 2 describes the fields used for the classification of the oral documents.

1. The oral archives in *Gra.fò*: from anthropology to linguistics to oral history

Oral archives are not peculiar to a single branch of knowledge. On the contrary, they appear as a virtual space in which different kinds of expertise convene and deal with unusual, original research questions concerning audio preservation, cataloguing, transcription, access rights management. The *Gra.fò* portal includes among others archives (and/or their subparts, i.e. collections) which originate from research projects conducted by linguists. Among others, the following archives have been digitised and catalogued:

1. **Alto Mugello**: interviews based on linguistic questionnaires and spontaneous speech, collected in the area of Firenzuola (province of Florence) during a research project leaded by Gabriella Giacomelli in the 80s.
2. **Atlante Lessicale Toscano (ALT)**: interviews based on linguistic questionnaires gathered from 1974 to 1986 by a group of interviewers, mostly graduands in Italian Dialectology, under the direction of Gabriella Giacomelli (Facoltà di Lettere, Università di Firenze). The field research covers the entire territory of the region (245 inquiries in total) on the basis of a questionnaire consisting of 745 questions, both onomasiological and semasiological.
3. **Silvia Calamai**: recordings of dialectological fieldwork conducted in 1995-1997 in the district of Leghorn (collection ‘Sassetta’), ethnotexts about wheat production collected in the area of Calenzano for the project “Atlante Generale dell’Alimentazione Mediterranea” (collection ‘AGAM’), interviews with the playwright Ugo Chiti (collection ‘Chiti’).
4. **Carta dei Dialetti Italiani (CDI)**: recordings collected in the 60s and 70s by scholars and students involved in a nation-wide research project on Italian dialects. Every region of Italy was at the time represented by several informants with fieldwork sessions recorded on tape. Today, these recordings are among the oldest available materials on Tuscan dialects [3].
5. **Seminari di Linguistica Generale**: investigations conducted since 2006 by university students who attended seminars or defended Master’s thesis in general linguistics at the Facoltà di Lettere e Filosofia di Arezzo/Dipartimento di Scienze della formazione, Scienze umane e della Comunicazione interculturale (Università degli Studi di Siena).

Many archives (and/or collections) contain material gathered in an anthropological, folkloric or ethnomusicological perspective:

1. **Edda Ardimanni**: folk tradition and folk music collected by different scholars (lullabies, *ottava rima*, narrative songs, *stornelli*, poems).

2. **Roberta Beccari (collection ‘Letteatura popolare’)**: recordings about folk literature and culture collected in the period 1986-1987.
3. **Vanna Brunetti**: university thesis research carried out with Gastone Venturelli between 1973 and 1975, concerning the dialect and folk traditions of Catena (a small village between Prato and Pistoia). This material includes folk songs, lullabies, jingles, proverbs, riddles, tongue twisters, fairy tales, legends, prayers, poems, and interviews about folk medicine, healing rituals, ancient formulas and incantations to cure or banish illness, both physical and mental.
4. **Anna Buonomini**: oral documents of ethnographical and folkloric interest collected in the 80s and 90s in the Montagna Pistoiese. The investigation was structured in three main research strands: historical toponomastic and folktales about the popular legendary figures of Roland and Annibale (collection ‘Cappel d’Orlando’), a research about the chestnut culture (collection ‘Castagno’), and memories of the Second World War (collection ‘Studenti’).
5. **Paolo De Simonis (collection ‘Canti popolari del Mugello’)**: university thesis research conducted in 1970-1971, concerning the dialect and folk traditions of the Mugello valley (Borgo San Lorenzo, Vicchio, Scarperia, Firenzuola, Marradi). The recordings include narrative songs, lullabies, jingles, proverbs, riddles, tongue twisters, folktales, poems, and interviews about ancient rural rites (e.g. the practice of carrying lighted torches around newly sown fields to promote fertility and ensure good crops, known as “far lume al grano”).
6. **FLOG**: the sound archive of the Federazione Lavoratori Officine Galileo (FLOG) is stored at the Mediateca delle Tradizioni Popolari di Firenze. It consists of a large amount of recordings collected during 30 years, concerning folk music and traditions (collection ‘Gilberto Giuntini’), legends, short stories, fairy tales from Castelnuovo Berardenga (collection ‘Nunzi Gioseffi’).
7. **Sergio Gargini**: research conducted during the 70s, 80s, and 90s by the ethnomusicologist Sergio Gargini with the collaboration of other members of the “Collettivo Folcloristico Montano”. The overriding interest of the fieldworker concerned the folk music and traditions of the area of the Montagna Pistoiese (narrative songs, *stornelli*, *contrasti in ottava rima*, *maggi* etc.). In addition, the archive conserves material of historical, ethnographic, and anthropological interest: interviews about living conditions of charcoal-burners, lumberjacks, peasants and labourers (mostly steelworkers employed at the Società Metallurgica Italiana in Campo Tizzoro), autobiographical memories of immigrants who moved from the Apennine’s villages just north of Pistoia to USA in the early 1900s.
8. **Benozzo Gianetti**: recordings of folk literature and music collected in the 80s in Valdera (Pisa). The research project was conducted in collaboration with teachers and schoolchildren from Ponsacco and Cascina Terme. The children interviewed their relatives asking about traditional songs, legends, little stories, jingles, and prayers. These oral documents were published in the book “Canti popolari della Valdera” (Gianetti 1984).
9. **Duse Lemetti - Gruppo Vegliatori**: research conducted in the 80s and 90s in order to preserve and transmit the oral tradition of Garfagnana (Lucca). The recordings were collected during evening get-togethers (“veglie”) in Galliciano.

10. **Museo del Bosco:** material collected in order to create an ethnographic museum of forest economy (“il Museo del bosco”, Sovicille, Siena).

Some archives provide rich resources for the study of contemporary history and sociology, especially for historians and sociologists interested in 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:

1. **FLOG (collection ‘Andrea Grifoni’):** life stories of labourers, memories of the labour movement and of the battles fought by the trade unions in the Officine Galileo di Firenze.
2. **ASMOS (Archivio Storico del Movimento Operaio e Democratico Senese):** memories of the labour movement in Siena, including political meetings, interviews, conferences, seminars, demonstrations, etc. The main topics are political activism, workers' struggles, anti-fascism, Italian resistance movement, racial persecutions, the Metayage system.
3. **Neri Binazzi (collection ‘Facchini’):** interviews with the workers of the Cooperativa fiorentina CFT (Cooperativa di Facchinaggio e Trasporto) collected since 1999.
4. **Cappelli di paglia:** memories and life stories collected in the 80s by Gianbruno Ravenni and Andrea Grifoni. The main topic of their research is the history of straw hats manufacturers, flourished in Signa and Lastra a Signa (Florence) from the late 18th century to the first half of the 20th century.
5. **Dina Dini:** research conducted in 1994-2005 by the Centro di Documentazione Storica della Civiltà Contadina “Dina Dini” in the South of Tuscany (Pieve Santo Stefano, Arezzo). The archive contains interviews on traditional peasant material and non-material culture (collection ‘Storia orale’), life stories of emigrants (collection ‘Emigranti’), and collective interviews concerning different aspects of the social life of the peasantry in Tuscany (collection ‘Circolo di studio’).
6. **Elba:** research conducted in 1999-2001 by the historian Tiziana Noce who interviewed more than 30 people from the Elba island (miners, sailors, peasants, labourers and employees from Rio Marina Capoliveri, Porto Azzurro and Portoferraio). The fieldwork was mainly focused on individual experiences and memories of the Second World War, the crisis of the mining industry, the experience of immigrants who moved from the Elba island to other countries (mainly USA and Australia), and the development of tourism in the second half of the 20th century.
7. **Roberto Segnini:** university thesis research conducted in 2008 about the textile industry in Prato in the 50s and 60s.

2. Classification of oral documents

A software tool for digital data cataloguing (AudioGRAFO) has been implemented within the *Gra.fo* project [4]. The tool allows the cataloguers to describe both the archives (and their subdivisions) and the single oral documents:

- information about the archives – name, place of conservation, existence of the ‘Tell something about your archive’ interview¹, privacy limits, description, motivation of the research, date of transfer to the project, owner.
- information about the oral documents – title, content, date and place of collection, data about the researcher and the speaker, available bibliography, accompanying materials, classification of the document, aims of the single recording, keynotes.

In *Gra.fo*, four criteria are used for the classification of oral documents, according to [5]:

1. Type – cataloguers distinguish between the following types:
 - controlled events (i.e., events under the researcher’s direct control, e.g. interviews) vs. uncontrolled events (e.g. documents collected with hidden recording modality);
 - sung documents (e.g. lullabies, *ottava rima*, *canto lirico monostrofico*) vs. spoken documents (e.g. interviews, riddles, fairy-tales);
 - formalised documents (e.g. lullabies, riddles) vs. non-formalised documents (for spoken documents, e.g. interviews and free speech) or improvised (for sung documents, e.g. *ottava rima*).

2. Topic – Cataloguers can chose among about 130 different topics (such as 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, Museography, Music festivals, Nazism, Peasant culture, Peasant traditions, Political history, Politics, Postwar 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, etc.). Since only one topic per document can be chosen, other relevant topics can be included in the keyword list.

¹ The owners of any archive which lacks a proper bibliography or accompanying materials are interviewed about the motivation and aims of their research. The interviews, called ‘Tell something about your archive’, offer the key for interpreting and cataloguing the archives, which are often clearly understandable only by the very researchers who collected them.

3. Genre – Cataloguers can choose among about 40 different genres (such as 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 etc.). Creating a fixed taxonomy for such an interdisciplinary project has proven really difficult, since the available taxonomies are partial (i.e. they refer to a single field of study) and often blur the boundaries between genres and topics.
4. Language variety – Cataloguers can chose among about 30 different varieties. Tuscan varieties have been divided into urban varieties (Arezzo, Florence, Grosseto, Leghorn, Lucca, Massa, Pisa, Pistoia, Prato, Siena), areas of influence and transition areas, according to the taxonomy proposed by Luciano Giannelli [6,7]. The sociolinguistic motivations for such a choice are twofold: first, cities vehicle linguistic identity and usually influence the surrounding areas; second, Tuscany does not have a hegemonic centre which can influence the whole territory of the region [8,9].

The creation of an archive incorporating the main oral archives of the region has involved different, interconnected stages of work. It has been necessary to lay the foundations for an interdisciplinary dialogue between linguistics, anthropology, computer and archival sciences. Such an effort has produced more than 2700 hours of digitised audio and more than 2200 catalogued oral documents at the time of writing this paper, while progressively more archives are offered to the project in order to be preserved.

Acknowledgment

This work is supported by *Grammo-foni. Le soffitte della voce (Gra.fo)* Project (Scuola Normale Superiore, Pisa & Università degli Studi di Siena, Dipartimento di Scienze della formazione, scienze umane e della comunicazione interculturale). PAR FAS 2007-2013 Regione Toscana Linea di Azione 1.1.a.3.

References:

- [1] Contini G., Martini A. 1993 *Verba manent. L'uso delle fonti orali per la storia contemporanea*, Roma: Nuova Italia Scientifica.
- [2] Andreini A., Clemente P. 2007 *I custodi delle voci. Archivi orali in Toscana: primo censimento*, Firenze: Regione Toscana.
- [3] Calamai S., Bertinetto P.M. 2012 "Per il recupero della Carta dei Dialetti Italiani", in T. Telmon, G. Raimondi, L. Revelli (eds.), *Coesistenza linguistiche nell'Italia pre- e postunitaria*. Atti del XLV Congresso internazionale di studi della Società di Linguistica Italiana (Aosta/Bard/Torino 26-28 settembre 2011), Roma: Bulzoni. 335-356.

- [4] Bressan F., Canazza S. 2013 A Systemic Approach to the Preservation of Audio Documents: Methodology and Software Tools, *Journal of Electrical and Computer Engineering*, Article ID 489515, doi: 10.1155/2013/489515, <http://www.hindawi.com/journals/jece/2013/489515/>.
- [5] Calamai S. 2011 Ordinare archivi sonori: il progetto *Gra.fo*, *Rivista Italiana di Dialettologia* 35. 135-164.
- [6] Giannelli L. 1988 “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. 594-606.
- [7] Giannelli L. 2000 [1976] *Toscana*, Pisa: Pacini.
- [8] Agostiniani L., Giannelli L. 1990 “Considerazioni per un’analisi del parlato toscano”, in M.A. Cortelazzo e A.M. Mioni (eds.), *L’italiano regionale*. Atti del XVIII congresso internazionale di studi della Società di Linguistica Italiana, Padova-Vicenza, 14-16.IX.1984, Roma: Bulzoni. 219-237.
- [9] Calamai S. 2011 “Toscana”, in R. Simone (ed.), *Enciclopedia dell’Italiano*, Roma: Istituto dell’Enciclopedia Italiana, II: 1506-1509, http://www.treccani.it/enciclopedia/dialetti-toscani_%28Enciclopedia_dell%27Italiano%29/.