Editorial

Why This Journal, Why Now

Silvia Calamai

University of Siena, Italy
E-mail: silvia.calamai@unisi.it

A tutte le voci che non abbiamo ancora ascoltato

Oral archives are artefacts at the crossroads of many fields of the social sciences and humanities, from linguistic disciplines (phonetics, dialectology, sociolinguistics, corpus linguistics) to oral history and ethnography, sociology, anthropology, and psychology, not to mention speech technologies (natural language processing and automatic speech recognition). Nevertheless, they are still "underrepresented in scholarship, teaching, and the public imaginary" (Clement 2024, xi). The educational potential of oral archives has also been exploited very little, and they appear to be an under-exploited resource in the context of intangible cultural heritage as well (with the exception of some ecomuseums networks – on which we will return later in this text).

There might be many reasons for this low utilisation. Firstly, one might recall that contemporary Western societies are mostly visual societies, and researchers from these countries dominate the scientific discourse in science, producing indirectly an initial bias toward vision (see Hutmacher 2019). But in addition to the dominance of the visual dimension, structural and contingent factors must also be considered, such as the fragmentation of archives, the lack of accessibility of institutional repositories and data centres, and the legal issues related to their reuse. These are all factors that undermine the very survival and reusability of oral sources. They are considered ephemeral and in the struggle for visibility – not only academic – they cannot compete with the written medium. Historically, and typically, archival records are texts, not voices. It is a fact that speech appears to be a severely underutilised material in the cultural heritage domain, thus representing a hidden treasure and an El Dorado of data.

Even the terminological labels are in some ways a reflection of such weakness and fragmentation. The words used in the single disciplines do not appear to be totally consistent. *Oral*, *speech* and *audio* may be used more or less indifferently with the meaning of a structured collection and repository of audio material; nevertheless the various labels refer to different disciplinary traditions. For instance, *speech archive* appears to be probably the most frequent label among linguists (together with *speech corpora*), while among oral historians and anthropologists, *oral archive* is much more widely used. The label *audio* – or even *sound – archive* usually covers a broader range, including music, too. Likewise, *archive*, *corpus* and *collection* are sometimes used as synonyms and sometimes in a mutually exclusive way according to the relative disciplinary practices. Within the editorial board of this publication, the choice fell on 'oral' because the term seemed to us more inclusive and less disciplinarily connoted (although a quantitative study on what scholars call the artefacts they produce is yet to come).

Oral Archives Journal aims to fill a gap in the Italian academic publishing scene, where journals dedicated to the world of archives mainly refer to 'written' sources. It wants to represent a virtual meeting place between the academic world, the various institutes of Ministry of Culture and its peripheral organs (Soprintendenze), and independent researchers as well. There have been journals dedicated to oral archives in the past. *Fonti orali* was a fundamental publication, directed by Luisa Passerini from 1981 to 1984, and by Daniele Jalla from 1985 to 1987, and published in Turin in 13 issues. The journal highlighted relevant attention to the methodological and ethical aspects of fieldwork. In the very same years, the early issues of *Rivista Italiana di dialettologia* also showed a non-episodic interest in data collection and transcription (together with the issues of *Bollettino dell'Atlante Linguistico Italiano*). Nonetheless, what is lacking today is an interdisciplinary point of reference connecting different fields of knowledge and, above all, bringing (very) different actors into the dialogue.

As outlined in the *About the Journal* section [https://riviste.fupress.net/index.php/oarj/about], the topics covered by the journal are related by two different areas: the life cycle of the archives, from production and curation to preservation and reuse; and cross-cutting dimensions of all the four aforementioned phases, with particular attention to the theoretical and methodological aspects of transcription, the technological tools, the legal and the ethical issues.

Special attention is devoted to the curation phase of the life cycle. In recent years we have been able to verify first-hand how there is also a lack of a publishing venue that can enhance the archival work of arranging and cataloguing oral sources. Taking care of one's sound records is a task that is often time-consuming and resource-intensive, but, on the other hand, is also a crucial step in ensuring the transmission and dissemination of their contents, as well as the replicability of studies on them. Exactly for this reason, *Oral Archives Journal* aims to put the spotlight on this activity usually carried out behind the scenes by the scholars of orality, in the firm belief that a well-structured oral archive should be considered a product of research. Adequate space is therefore given to presentational essays of oral archives, based on the evaluation of the archival arrangement of the materials themselves. *Oral Archives Journal* advocates a change in mentality for a successful transition towards open science in order to make data available to the scientific community and reusable by others. Yet this transition cannot be achieved by single individuals without institutional support at various levels (single university/research centre; national evaluation agencies; national and international funding

agencies). A partially similar position is outlined in Garellek et al. (2020, 1), where the difficulty of evaluating "data curation" activity in terms of academic careers is pointed out, too, as an aspect that is complicated when considering different disciplinary fields, their traditions and styles. Among phoneticians and phonologists, for example, there is a clear awareness of the need to proceed according to a change of mindset as well as a lack of a unified set of guidelines, since the efforts to curate and archive data sets totally depend on the goodwill of individuals: "a collective turn to make data curation and archiving mandatory will be highly beneficial for the phonetic sciences" (Garellek et al. 2020, 2; see also Cho 2021; Didirková and Simon 2024). From our perspective, both curating and archiving data become therefore an integral part of publishing. On the opposite side (i.e., the user, not only the producer), Clement (2024, xii) also emphasises this point: "If you are able to listen to archival recordings, it might entail hundreds of hours of exploration and close listening, because the metadata that librarians and archivists create is often based on the limited information they find on old, incomplete, confusing, or wrong and unverified labels". This is why we dare to believe that a journal with these aims is today necessary. Exactly for this reason, the composition of the Editorial Board [https://riviste.fupress.net/index.php/oarj/about/editorialTeam] – which also includes many young scholars – is diversified and represents many different disciplines, just as different are those represented by the Advisory Board [https://riviste.fupress.net/index.php/oarj/about/editorialTeam] for the five-year period 2023-2027. To the journal's Editorial Board and to the Advisory Board goes my heartfelt thanks for agreeing to be part of a new and in some ways unusual undertaking, on the Italian scene. Thanks are due also to the fourteen reviewers involved in the evaluation of the articles: all their names will be made public at the end of the first three years of the journal's life, along with the reviewers of issues 2 and 3. The index and the content of this first issue bear witness to this wealth – where dialectologists, sociolinguists, computational linguists, engineers, physicists and archivists appear. Not only from academy but also from the institutes of the Ministry of Culture entrusted with the preservation of these precious vocal treasures.

The path that led to this publishing project is very long and comes from afar. In the following sections I will try to tell its story, the fellow travellers, the obstacles, and the challenges ahead.

AN AUTOBIOGRAPHY THAT IS NOT JUST ABOUT ME

Recorded voices and oral archives have accompanied my own life, starting from the age of 22 – first as a student, then as a scholar. I did my very first dialectological interview when I was a university student, in the countryside where my father was born. Since then, I never stopped recording voices.

I wrote part of my thesis in a room at the University of Florence where hundreds of old reels and accompanying materials were stored. I stared at them wondering about who could listen to them again. They came from an extraordinary geo-linguistic project carried out during the late post-war period in Italy, called the *Carta dei Dialetti Italiani* (CDI). The CDI project provided a detailed picture of the dialects of Italy that existed in the early 20th century, which were captured through interviews with middle-aged or elderly people at the time. The archive collection holds approximately 1,100 open reels – accompanied by written

reports – corresponding to nearly 1,600 hours of recorded speech. The CDI archive was a unique treasure that remained sorely inaccessible to the scientific community. The enterprise – conceived and directed by Oronzo Parlangèli (University of Bari) – was originally designed as a *questionnaire* made by 529 entries, together with the collection of a dialectal version of the *Parable of the Prodigal Son* from every Italian municipality (Parlangèli 1972). Today, it still represents one of the most important projects of dialectal material collection of the last post-war period. Moreover, the project was characterised by the systematic use of the tape recorder and by the determination to preserve the oral material collected. After Parlangèli's unexpected death, the CDI research centre, established by the National Council of Research (CNR), moved from Bari to Padua under the direction of Manlio Cortelazzo. Unfortunately, the materials were then dispersed between the CNR's Padua Centre of Dialectology, the University of Bari, the Colombaria Academy in Florence, and to some of the researchers who took part in the project. The recorded materials were at risk of deterioration, and it was crucial that someone intervene before it was too late. An 'official' inventory of all the sound and paper materials of the CDI has not yet been achieved. The only inventory updated to 2011 was conducted by myself and my colleague Pier Marco Bertinetto (Calamai and Bertinetto 2012). With the CNR's consent, we gathered together, in the Laboratorio di Linguistica "Giovanni Nencioni", all of the materials scattered throughout the Italian peninsula. Our goal was to preserve and disseminate this material, while stimulating reflection on the importance of returning publicly funded research to the citizens themselves. The CDI audio reels had been forgotten for almost 40 years by the Italian scientific community, despite the fact that the enterprise had received public funding.

That room also hosted the staff of another geo-linguistic enterprise of the past century: the *Atlante Lessicale Toscano*. It was conceived and promoted at the beginning of the 70s by Gabriella Giacomelli, professor of Italian Dialectology at the Facoltà di Lettere e Filosofia of the University of Florence. The field investigations (245 as a total, including the supplementary investigations and those which were repeated) were carried out between 1974 and 1986 and were based on a questionnaire made up of 745 questions. The researchers (mainly graduate students/PhD students working with Gabriella Giacomelli) interviewed 2193 participants, who were chosen in every locality according to sociolinguistic factors such as age, sex, social condition, and education, in order to document variation and ongoing changes within the same community and the serious loss of dialectal lexis among younger speakers, as well as the variation in space. The 224 localities under investigation were mainly small agricultural settlements within the regional administrative area, including non-Tuscan speaking communities such as 'Lunigiana' and 'Romagna Toscana'. After completing the field research, the idea of creating a paper linguistic atlas following the examples of the *Sprach- und Sachatlas Italiens und der Südschweiz* (AIS, now entirely digitised and accessible via the following url: https://navigais-web.pd.istc.cnr.it) and the *Atlante Linguistico Italiano* (ALI: https://www.atlantelinguistico.it) was abandoned, and all the collected material was stored in a database which is still accessible online via ALT-WEB, under the direction of the Istituto di Linguistica Computazionale Antonio Zampolli (ILC) of the CNR of Pisa (https://dbtvm1.ilc.cnr.it/altweb/RT_ALT-WEB_home.htm). While the database (BD-ALT) contains the corpus of all the answers, the recordings, which the researchers were not required to make, are in some cases incomplete, as they are related only to some par-

ticipants and some parts of the questionnaire. In addition, the researchers bought the tapes themselves and were not obliged to give them to the research group.

I went into that room many times in order to consult different theses and reports and to leaf through the pages of the fieldwork memoranda taken in the 1960s and 1970s in order to glean their secrets. It was somehow easy to access theses in Dialectology (carried out by dozens and dozens of students under the supervision of Gabriella Giacomelli), avoiding all the complicated official routine that the University library demanded in order to consult unpublished dissertations. In that room everything could be read as long as it was properly and correctly cited – Gabriella Giacomelli thus cautioned. And that's what we did. It was somehow necessary to meditate on those recorded voices in a room filled to the brim with voices that could not easily be heard again at the moment. I was considering the fieldwork carried out by others, and at the same time I began my own fieldwork on the vernacular speech of a very small community in Maremma, Tuscany. I observed, interviewed, collected, and recorded hours and hours of speech, transcribing the most relevant parts myself.

At that time (it was about the end of the Nineties), I met several scholars who believed that transcriptions were much more important than the recordings themselves; many of them simply threw away their reels and compact cassettes once a transcription was completed. Analysis was deemed more essential than unrefined data. As one distinguished scholar told me sharply during a discussion: "I have to remind you that it is one thing to analyse data and another to just collect". The case for the importance of preservation, reuse, reproducibility and accessibility of oral archives had not yet been made. The destiny of all the recordings collected in the field by linguists and other scholars – thousands of hours of speech recordings gathered for different purposes (e.g., linguistics, anthropology, social sciences, oral history, ethnomusicology, …) – was uncertain. When I discussed my PhD thesis one of the members of the committee asked me where I would store all the audio materials I had collected for the analysis of the Western Tuscan vowel systems. I was totally unaware of that issue and I did not know how to answer.

Then came the time of corpora creation in the Laboratorio di Linguistica "Giovanni Nencioni", Pisa. As a post-doc researcher, under Pier Marco Bertinetto's guidance, I had the opportunity to participate in national projects that were beginning to build oral resources for the Italian language (*Archivio di Varietà di Italiano* Parlato – AVIP, *Archivio del parlato italiano* – API, *Corpora e lessici dell'italiano parlato e scritto* – CLIPS; see Albano Leoni 2007) – that is, something that should have been investigated by others. And eventually the funded projects period began. In 2011, the Region of Tuscany provided financial support to Pier Marco Bertinetto and me for a two-year project called *Grammo-foni. Le soffitte della voce (Gra.fo)*, to create a digital archive combining the main oral archives of the region. Since the development of the UNESCO Convention for the Safeguarding of Intangible Cultural Heritage (2003) and the widespread use of the term "intangible cultural heritage", the importance of digitisation, cataloguing and dissemination of research has become clearer and more comfortable to many. Yet uncertainty endures and some scholars still preferred to keep their cassettes safe in their drawer and their audio files in their personal data storage. Tuscany was a privileged area for working on oral documents, as it abounds with both public and private oral archives, collected in different fields of research by both scholars and amateurs. The majority of these archives are analogue and therefore risk deterioration unless they are trans-

ferred to the digital domain. This project was jointly conducted by the Scuola Normale Superiore and the University of Siena (Regione Toscana PAR FAS 2007-13). Although we conceived of the project as linguists, we immediately realised how cross-disciplinary it was, and that we needed help from several other colleagues. It has been necessary to lay the foundations for an interdisciplinary dialogue between archival science, linguistics, law, anthropology, and computer sciences. The *Gra.fo* project has undertaken the challenging task of gathering different kinds of expertise and building a digitisation and cataloguing system with the aim of creating a regional network for the management of oral archives. With over 2,800 hours of digitised recordings and an incredibly vast range of document types (oral biographies, linguistic questionnaires, folk songs, questionnaires, lullabies, etc.) and topics covered (history, politics, economy, anthropology, linguistics, folk literature, etc.), the *Gra.fo* archive was considered a unique and exemplary accomplishment in the Italian panorama (Calamai and Bertinetto 2014; Calamai, Ginouvès and Bertinetto 2016).

But more than the enterprise itself, the relationships that this enterprise enabled to weave counted. Not only was the world of Italian linguists contacted. It was Pier Marco Bertinetto, as the principal investigator of the project, who wove the threads of the correspondence exchanges. He wrote to several colleagues and asked for an account of their tapes, whether they had kept surveys of the *Carta dei dialetti italiani*, or if they had tapes to have digitised. That private correspondence (in which I was cc'd, as co-coordinator of the project) is itself a relevant testimony to the subjective perception (by the insiders) of such requests: enthusiasm (*bravo! It's about time someone took care of it*), indifference (*I don't know where they are, I've lost everything*), annoyance (*with all the problems there are in the world, are you now asking for tapes?*). Instead, my 'political' task was to make connections with those who had travelled the road well before us. I invited Giovanni Contini Bonacossi from the *Soprintendenza per i beni archivistici e bibliografici della Toscana* to see the digitisation laboratory that was set up in the Laboratorio di Linguistica "Giovanni Nencioni", thanks to the advice of Sergio Canazza. Which Giovanni Contini promptly did. And even in later years he would always mention to me the incubator for 'cooking' the tapes, which evidently impressed him. Then there was Pietro Clemente, who at that time worked at the University of Florence, and who had recently finished the Tuscan census on oral archives (Andreini, Clemente 2007). That reading opened up so many avenues and dramatically showed how much work there was to be done. The censuses that came after usually use that work as a starting point (as Piccardi, this issue, mentioned). I went to him as well and he put me in touch with Véronique Ginouvès from the *Phonothèque de la Maison méditerranéenne de sciences de l'homme* (Aix-en-Provence), and with the French oral archives world – who started asking questions about ethics and oral archives, from collection to their use and reuse (Ginouvès, Biliotti and Calamai 2018). I was then invited to join the editorial board of *Sonorité. Bulletin de l'Association Française des Archives Sonores, Orales & Audiovisuelles*. This journal has also been a source of inspiration for us.

Then I got in touch with the world of oral historians. In 2015, the Italian Association of Oral History (AISO) organised a conference in Trento, November 2015 (*Buone pratiche di storia orale. Questioni etiche, deontologiche, giuridiche*: Bonomo, Casellato, and Garruccio 2016; Casellato 2021) to present the document *Good practices for oral history* (https://www.aisoitalia.org/buone-pratiche), in which the legal and ethical issues were discussed. Moreo-

ver, the conservation of oral sources was the topic of a round table, which involved scholars of different disciplines and representatives of different institutes from the Ministry of Culture. In addition, it served to bring out the many hurdles that would have to be faced, and it also made it possible to establish a preliminary relationship between the Italian Association of Oral History and the Italian Association of Speech Sciences (AISV), which I represented there. Precisely from 2015 started my close dialogue with Alessandro Casellato and the AISO community. In the same year I got in touch, through Francesca Frontini and Monica Monachini, with the European infrastructure of CLARIN – Common Language Resources and Technology Infrastructure. Together with Stef Scagliola, Henk van den Heuvel, Arjan van Essen, and Christoph Draxler, the *Speech and Tech group* was founded, and several workshops in Europe have been organised since 2016 (the detailed list is on the following website: https://speechandtech.eu/workshops), trying to create stronger relationships and a common lexicon between humanities scholars and speech technologists. The joint work with legacy data culminated in the CLARIN project *Voices from Ravensbrück*, in which Fabio Ardolino also participated (Calamai et al. 2022).

The time was ripe to think about pulling the strings together of this dialogue, laboriously created in those years. A round table on *Sound archives at the crossroads of phonetic sciences, humanities computing and digital heritage* (15th AISV National congress, Arezzo, February 2019) strengthen the partnership between AISV and AISO, and another demanding path started. A detailed account of the event can be read in the *Introduction* to the publication edited by Piccardi, Ardolino and Calamai (2020), which also contains paper versions of some of the speeches held at the round table. The participants came from very different research fields and had very different roles, both at universities, research centres and institutes of the Ministry of Culture. In that context, on the impetus of AISV, AISO and the *Soprintendenza per i beni archivistici e bibliografici della Toscana* of Tuscany, a national working table – the so called *Tavolo permanente per le fonti orali* – was set up aimed at defining a *Vademecum* for the conservation, description, use and reuse of oral sources. The group also consists of representatives from the main institutes of the Ministry of Culture working on descriptive practices and standardisation of document processing (Istituto Centrale per i Beni Sonori e Audiovisivi, ICBSA; Istituto Centrale per il Catalogo Unico, ICCU; Istituto Centrale per il Catalogo e la Documentazione, ICCD; Istituto Centrale per gli Archivi, ICAR), together with representatives of the institutes headed by the National Institute Ferruccio Parri, of the Soprintendenza archivistica del Piemonte, of the CLARIN-IT infrastructure and of the Computational Sonology Centre of the University of Padua. The purpose of the *Tavolo permanente per le fonti orali* was to make available to individual researchers and research groups some clear and essential indications on how to guarantee the conservation of oral archives, their correct archival description, and finally their use and reuse. The *Vademecum per il trattamento delle fonti orali* (*Vademecum for the treatment of oral sources*) is therefore the culmination of a participatory work initiated in 2019 by the *Tavolo permanente per le fonti orali*, which involved linguists, oral historians, archivists, speech technologists from various institutions. In various meetings, which took place throughout 2019 and 2020, and for the first time in Italy, different subjects found themselves dialoguing and working together, trying to build a common lexicon to foster a conscious use of oral data, i.e., of all the audio material collected in the past and to be collected in the future. The resulting *Vademecum per le fonti orali*

was published in 2023 as a monographic issue of the "Quaderni della Rassegna degli Archivi di Stato" (no. 114), and is designed for the researcher working on oral resources, even on an occasional and independent basis, for institutes that find themselves dealing with heterogeneous and insufficiently codified material, and for disciplinary associations whose members make use of oral data. The working group has produced guidelines dedicated to the topics of preservation, description, use and reuse, with a series of appendices containing up-to-date lists of preservation and digitisation institutions, examples of descriptive forms, examples of agreements and informed consents, which can be adapted to the specific research context. A section on the different forms of transcription of oral resources was added in 2023 (edited by Chiara Celata).

The *Tavolo permanente per le fonti orali* disseminates the guidelines contained in the *Vademecum*. It monitors the research activities of institutions and individuals working in the field of oral archives, in relation to technological tools, procedural standards, the current legal framework in Italy and Europe, and publications; it promotes the deposit of currently dispersed material in organised archives that can guarantee its preservation and access; and it promotes meeting opportunities by organising conferences and sponsoring training activities within the various disciplinary associations (above all, the Italian Association of Speech Sciences and the Italian Association of Oral History). The *Tavolo permanente per le fonti orali* has always been open to the contribution of young researchers. It has proven to be truly intergenerational, with scholars and enthusiasts of different profiles and at different stages of career (refer to the recordings of the annual meetings on the website of the *Tavolo*: https://sites.google.com/view/tavolopermanenteperlefontioral/il-vademecum-per-il-trattamento-delle-fonti-orali).

In the same years, oral archives of the past were reused in research projects in collaboration with Ecomuseums and other public bodies (see the *Landscape in Sounds through Eco-Museums network* LISTEN POR FSE 2014-2020 project, together with Rosalba Nodari, and now the Interreg CASTOUR project with the province of Lucca: https://interreg-marittimo.eu/web/castour/progetto). Again in the same years, Siena University, the Soprintendenza archivistica e bibliografica della Toscana and ILC-CNR tried to lay the foundations for an infrastructure dedicated to oral archives, *Archivio Vi.Vo.*, which has not yet come to fruition. There was a favourable situation, because the *Tavolo permanente per le fonti orali* was working in parallel and the project was going down unprecedented paths, thanks also to the contribution of Maria Francesca Stamuli at the Soprintendenza archivistica e bibliografica della Toscana, who combined in her skills both a doctorate in linguistics with Rosanna Sornicola and archival knowledge. *Archivio Vi.Vo.* aimed at exploring theory, methodology, and services for long-term preservation and secure access for audiovisual documents (Calamai, Piccardi, Pretto, Candeo, Stamuli and Monachini 2022; Valentini, Piccardi, Calamai and Stamuli 2023). In the initial stage, the services for long-term preservation and access were developed within the framework of CLARIN-IT, the Italian counterpart of CLARIN-ERIC. The archival model represents the final outcome of a journey from *Gra.fo* experience to the present day and was intended to be disseminated in the scientific community interested in managing oral records. Linguists, archivists, and engineers worked in the design of the model. The digital infrastructure realising the model should now be put in place through a PRIN project (*Roads to Oral Archives Development and Sustainability* ROADS, PRIN 2022: https://csc.dei.unipd.it/

roads-project/), and through the H2IOSC project (https://www.h2iosc.cnr.it/). It should also accommodate some of the previous projects we have mentioned on these pages, such as the *Gra.fo* project. The idea of transferring *Gra.fo* to other, newer infrastructures was also piloted during the *Gra.fo Reloaded* project, which started in 2022 and is extensively described in the paper by Duccio Piccardi in this issue. Through his support, during *Gra.fo Reloaded*, much of the groundwork for this journal was laid, from surveys of contemporary similar initiatives to an initial envisionment of its structure and role in the research communities of interest (Piccardi and Calamai forthcoming). At the same time, the *Gra.fo* portal was reactivated by the Scuola Normale Superiore, but only as an example of what was accomplished and available at that period, certainly not open to new accessions (i.e., new archives) or new registrations (i.e., new users). The profound changes that have taken place in the European and Italian legal framework since the birth of *Gra.fo* have prompted the adoption of a principle of caution (see the detailed explanation at https://grafo.sns.it/site/about).

In the meanwhile, the informal working group known as the *Tavolo permanente per le fonti orali* meets every year, on the Unesco Day for Audiovisual Heritage, to take stock of the situation: where we have got to, current activities, and future actions to be planned and organised. It is not easy to keep track of the archipelago of initiatives springing up in many areas, even in the same country. That is why the annual meetings of the *Tavolo* are so important. They represent a place where different people reason together on what to do. Specifically in the case of oral archives, more than in other human artefacts, Lorenzo Milani's invitation in *Letter to a Teacher* fits well, *mutatis mutandis*: "To come out of them together is good politics. To come out alone is stinginess" (section *Politics or stinginess*) (Scuola di Barbiana 1967, 1970). And precisely the last section of this introduction tries to outline the ethics and politics of oral archives.

ORAL ARCHIVES MIRRORING OUR TIME: THE CHALLENGES AHEAD

There are at least three domains that we need to be clear about when handling oral archives from a perspective that is not limited to the individual scholar and the individual research. The themes of long-term preservation and sustainability, those of serendipity and reuse, and those related to the ethical and legal issues are all cross-cutting and ubiquitous, and strictly intertwined. Each concept would require a separate discussion, which would not even be sufficient.

Let us begin with the issue of long-term preservation and sustainability. Both topics are extremely timely and, in some ways, risky, given that with the *National Recovery and Resilience Plan* (PNRR) funding many have thrown themselves into the digitisation business. This brings to mind the issues associated to *The afterlife of data* by Őhman (2024): his reflections were originally conceived for online life, but *mutatis mutandis* can also be applied to the world of oral archives. Since data do not die with the person who left it, to whom do the data of the dead belong? And crucially, what obligations do they impose on the living, especially when it emerges that it is not possible, after a 'diligent search', to find all the holders of rights? Casellato and Calamai (in press) mention historian Manlio Calegari's provocation, when he recalled how important contextual and accompanying documents are for making oral sources comprehensible; this was during a public meeting organised by AISO in 2013. He then posed

a radical question to his younger colleagues: can recorded interviews – in his case, 201 audio-tapes collected over twenty years with 73 people, some of whom he met several times, on the history of Genoese partisanship – really be understood by those who listen to them without knowing everything that interviewee and interviewer had in common and precisely what they did not say? Manlio Calegari wonders, now that he is old, what to do with those audio cassettes that have remained in his library until now:

> Destroy them or hand them over to someone, some institution that deals with such things? Herein lies the problem: I am sure – for real – of their uselessness to anyone who has not collected, heard and transcribed them all. Beyond the unsaid and the unwritten – which is no small thing – there is a mass of details collected, compared, clarified, integrated, and edited during ten years of work that have made them objects other than how they are presented. I imagine that those who have had such experiences know what I am talking about (my translation).

Should everything be preserved and maintained according to international standards? What to do with all the digital projects, more or less substantial, already closed and accounted for, that soon will become or have already become obsolete? Digital-born artefacts that originate and circulate across a range of formats and platforms are rapidly expanding and raising new challenges for research, archiving and collecting communities. Collecting, accessing and sharing digital-born oral archives presents a range of complex technical, legal and ethical challenges that, if unaddressed, threaten the archival and research futures of materials and records of the 21st century. Crucially, who decides what to preserve? Oral archives – both analogue and digital-born – raise deeply topical issues. Data are not archives themselves, but they can become archives (see also Cialone, Ginouvès, this issue). At some point someone who has or finds a drawer full of tapes realises that he or she does not have an indistinct mass of tapes but a real archive (it happens). The problem is that which is apparently meaningless now could acquire meaning with research questions that no one has thought of at the moment.

Difficult and uncomfortable decisions that no one has yet thought of stem from the tension between preserving 'everything' (since in 'everything' also lies that particular document now considered totally insignificant but perhaps a harbinger of future developments), or 'choosing' instead what to preserve because of the costs of digital preservation. Are the individual disciplinary associations that determine, hopefully in a participatory process, what to save and how? Yet associations and disciplines are somehow a 'tame plague' (*peste mansueta*), as the poet (and academic) Eduardo Sanguineti admonishes us in *Laborintus* section 11, vv. 1-2 (*la nostra sapienza tollera tutte le guerre / tollera la peste mansueta delle discipline*), and they carry with them biases of no small importance if they themselves were to lead the sorting. We are thus facing a more thorough selection. I will try to explain the reasons. Although oral archives are essential for better understanding 'underdocumented' communities (immigrants, women, the uneducated, the impaired, …), they are a 'representation' of the real world, which is the outcome of several choices (which type of people to interview, according to which parameters, on which phenomena). "A particular technical and sociopolitical context influences how and why artifacts resonate as culturally significant", Clements (2024, 7) admonishes. From this point of view, long-term preservation of oral archives would be a *selection of representation* (which archives to preserve, and which to dismiss). Representation is a necessary but not sufficient condition: it may be the case that certain preserva-

tion policies strongly reduce the representation of different voices, different subjects, and different communities, thus greatly reducing our knowledge of what is left unsaid, unrecorded, and finally unsafeguarded. We are thus moving into new, unknown territory; an ethics and a practice are all to be imagined and constructed. This journal would like to host such reflections, which open up previously unexplored knowledge and horizons.

To new horizons we also add a different look at old horizons, which are only apparently already known. Reusing existing data not only contributes to the advancement of knowledge but also offers a more economical approach compared to the collection of new data, which often requires significant investments of time, effort, and resources. Scholars in the social sciences and humanities have increasingly recognised the benefits of reusing and repurposing legacy data (Bornat 2003; Rogers and McAllister 2014). Exactly here serendipity comes into play. Serendipity applies to human beings as a whole ("we are a serendipitous species; many times in our evolution we have discovered something advantageous by going in search of something else", Pievani 2018, 79, my translation), and to oral archives in a special way. We cannot neglect the revolutionary significance of some rediscoveries in oral archives which from certain aspects can be considered a case of exaptation – to continue the biological metaphor. A limited list of them from the field of linguistics is, in our view, effective enough. Guy Bailey's discovery of the mechanical recordings of interviews with former slaves born between 1844 and 1861 in the Library of Congress and the subsequent publication of the transcripts has made it possible to explore this crucial source of data for the study of Black Vernacular English and provide direct evidence on its earlier stages (Bailey, Maynor, and Cukor-Avila 1991). Elizabeth Gordon's discovery of the Radio New Zealand Mobility Unit Recordings containing oral histories between 1946 and 1948 from the voices of speakers born between 1851 and 1904 (Gordon et al. 2004) led to the design of a sociolinguistic project on the Origin of New Zealand English (ONZE). The sociophonetic analysis of plosive consonants in the ONZE corpus gave rise to seminal research on the interaction between word frequency, repetition, conversation topics and speaker's age according to exemplar theory (Hay and Foulkes 2016). But even in Italy there are now various research groups pursuing these challenges (e.g., Nodari and Calamai, 2021; the PRIN PNRR 2022 *Space and time in oral archives* STOA, https://sites.google.com/unisi.it/stoaproject/home). Coming across archives of the past means interrogating them with the eye of the present and doing new research. Linguists reuse and repurpose oral archives collected by historians and anthropologists, while speech technologists reuse and repurpose legacy data in general (see Coro, Cutugno, Schettino, Tanda, Vietti, Vitale, this issue).

This is where ethics and law come in and question us as scholars and as human beings. One should not forget that oral archives are not monoliths; they are the outcome of a plurality of approaches and disciplinary traditions that do not always find points of contact. We can try to summarise two opposite traditions with Alessandro Casellato's words. On the one hand, there is a "predominantly artisanal" mode among Italian oral historians and linguists (especially in the field of dialectology) which Casellato (2021, 13) summarises in these terms: "fiducial and informal relationship between the researcher and his sources, often in-house management of the entire source production chain, in-depth interpretative work conducted on small data or even single witnesses"; on the one hand, there is a "Fordist-taylorist" model

of US matrix with "large projects of massive collection and archiving of oral sources, which have promoted forms of specialisation and division of labour and even sub-contracting, con-tractualisation of the relationship with the interviewee, prevalence of collection over inter-pretation, recourse to information technology for uniform cataloguing and dissemination of products through mainstream and commercial channels" (my translations).

Given such dissimilar, and in some ways opposing, frameworks, should one protect the voices to the highest degree, or spread them with extreme caution, or spread them with some caution, or even spread them without caution? Disseminating records of the past on the inter-net puts a strain on intergenerational responsibility regarding promises made to interviewees who may now be deceased (Garruccio 2021, 119). Actually, the issue can also be seen from a different, in some ways opposite point of view, as some linguists argue: "Anonymizing data, or keeping the data to oneself, may not be as ethical as it seems. One needs to ask to what extent the real motivation is to protect the interviewees (the speakers, the informants – every discipline has its own labels), and to what extent the motivation is self-protection (to be safe-guarded against potential trouble). Lack of data conservation and of access to the data could also be seen as unethical behaviour, to the extent that it deprives the community of a useful resource" (Garellek et al. 2020, 6-7). Since the contents of archives are varied and diverse, we have a great need for case and feasibility studies, in order to understand, from time to time, how it is best to act. It is a matter of making the transition from a private choice to one that is public in its decision-making flow (even if the data, in the end, may not be entirely open), and we strongly believe in the role that this *Journal* might have in the public debate. Such is the case with Anna Maria Bruzzone's oral archive containing the voices of the patients from the Arezzo neuropsychiatric hospital, a serendipitous find on which many among us from the Edi-torial Board grew up (Calamai, Kolletzek, and Kelli 2017). It happened that the person enti-tled did not allow the reuse of the data, despite all our assurances. It happened also that there were people totally unaware of having had a relative in a psychiatric hospital, whose voice was recorded by Anna Maria Bruzzone. In this case, they were not informed of the finding: that is, considering the pros and cons of the situation, it was deemed appropriate not to change their knowledge of the family history. In some particular cases it was better left unsaid.

Given all these reasons, we need a reflection that might be, in its final stage, truly trans-formative. The care of research archives and the ethical and legal chain that these data require must not be considered as 'archive stuffing' and 'consent washing' respectively, but a healthy and community-friendly way of doing research. This is why it is essential to train ear-ly stage researchers who are truly interdisciplinary, who do not experience the legal aspects as a bureaucratic task to be fulfilled, a box to be ticked. It is true that we are witnessing the growing complexity of the legal system, with a multiplicity of centres and levels of regula-tory production, from local authorities to the Regions, from the State to the European Union and even to independent authorities. It is also true that researchers must cope with a rapidly changing reality, which is becoming increasingly complex, since interests, values and rights often conflict with each other and are very difficult to balance in the formation of legal rules (Tafani 2023, 76-77). But those who work with oral archives cannot look at the law as a fear-some threat, from which we must defend ourselves, as a 'risk'.

Prodi (2020) has well depicted the hardening of every aspect of everyday life into a legis-lative and judicial regulation that extends day by day into areas where previously only ethics

and morality acted. In our specific case, this makes the work an ongoing extenuating negotiation with all the actors involved in the life cycle of oral archives – a negotiation that in some ways risks alienating the younger generations of researchers from this artefact. The legislation on privacy and copyright, which in theory should aim to defend and protect the citizen (and hopefully the researcher), combined with the development of information and communication technology, only acts as a multiplier effect of a kind of paralysis, a further engorgement of social life and even a crisis of individual responsibility. It is incomparably easier and more peaceful to work on written sources. Yet we still do want to work with oral archives. One of the aims of this *Journal* is therefore to try to show the other face of the law. As Zagrebelsky (2007, 58) puts it, the one that inevitably needs the humanities (the other "sciences of culture", in Zagrebelsky's words: p. 156). I mean the law that does not produce norms and ties, closed within the constraints of mere "legal technique", but the law that takes due account of the weight and risks of creativity: "weights and risks that, at certain times and in certain cases, can turn consciences upside down" (Zagrebelsky 2007, 156, my translation). This is especially true of oral archives from the past, because of their contents and what they call into question: the relationship between the interviewer and the interviewee, the relationship among the heirs, the persons entitled, the rightful owners, and the choice to make the contents or part of the contents public. The human relationships they contain and which can put narratives, feelings, knowledge back into circulation, as well as every single choice the researcher has to make in the life cycle of archives, they all urge us to repeat: *Handle with care* – as a famous title on oral sources says (Bonomo, Casellato, and Garruccio 2016).

Proceed slowly. Listen.

REFERENCES

Albano Leoni, Federico. 2007. "Un frammento di storia recente della ricerca (linguistica) italiana. Il corpus CLIPS". *Bollettino d'Italianistica* (n.s.), 4: 122-30.

Andreini, Alessandro, and Pietro Clemente, eds. 2007. *I custodi delle voci. Archivi orali in Toscana: primo censimento*. Florence: Regione Toscana.

Bailey, Guy, Natalie Maynor, and Patricia Cukor-Avila (eds.). 1991. *The emergence of Black English: Text and commentary*. John Benjamins.

Bonomo, Bruno, Alessandro Casellato, and Roberta Garruccio. 2016. "«Maneggiare con cura». Un rapporto sulla redazione delle Buone pratiche per la storia orale". *Il mestiere di storico*, 8(2): 6-21.

Bornat, Joanna. 2003. "A second take: Revisiting interviews with a different purpose". *Oral History,* 31(1): 47-53.

Calamai, Silvia, and Bertinetto Pier Marco. 2012. "Per il recupero della Carta dei Dialetti Italiani". In *Coesistenze 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), edited by Tullio Telmon, Gianmario Raimondi, Luisa Revelli, 2: 335-56. Rome: Bulzoni.

Calamai, Silvia, and Bertinetto Pier Marco. 2014. *Le soffitte della voce. Il progetto Grammo-foni*, Manziana, Vecchiarelli.

Calamai, Silvia, Véronique Ginouvès, and Pier Marco Bertinetto. 2016. "Sound archives accessibility". In *Cultural Heritage in a Changing World*, edited by Karol Jan Borowiecki, Neil Forbes and Antonella Fresa, 37-54. SpringerOpen. https://doi.org/10.1007/978-3-319-29544-2_3.

Calamai Silvia, Chiara Kolletzek, and Aleksei Kelli. 2018. "Towards a protocol for the curation and dissemination of vulnerable people archives". In *Selected papers from the CLARIN Annual Confe-*

rence 2018, Pisa, 8-10 October 2018, edited by Inguna Skadina, Maria Eskevich, 28-38. Linköping: Linköping University Electronic Press.

Calamai, Silvia, Stef Scagliola, Fabio Ardolino, Christopher Draxler, Arjan van Hessen, and Henk van den Heuvel. 2022. "Ravensbrück interviews: how to curate legacy data to make it CLARIN compliant". In *Selected papers from the CLARIN Annual Conference 2021*, edited by Monica Monachini and Maria Eskevich, 1-9. Linköping: Linköping Electronic Conference Proceedings. https://doi.org/10.3384/ecp1891.

Calamai, Silvia, Duccio Piccardi, Niccolò Pretto, Giovanni Candeo, Maria Francesca Stamuli, and Monica Monachini. 2022. "Not Just Paper: Enhancement of Archive Cultural Heritage". In *CLARIN. The Infrastructure for Language Resource*, edited by Darja Fišer, and Andreas Witt, 647-65. Berlin: De Gruyter. https://doi.org/10.1515/9783110767377-025.

Casellato, Alessandro. 2021. "Introduzione. Buone pratiche per la storia orale: un cantiere aperto". In *Buone pratiche per la storia orale. Guida all'uso*, edited by Alessandro Casellato, 9-65, Florence: Editpress.

Casellato, Alessandro, and Silvia Calamai (in press). Gli archivi della ricerca. *La ricerca folklorica.*

Cho, Taehong. 2021. "Where we are at: Impact, special collections, open science and registered report at the Journal of Phonetics". *Journal of Phonetics*, 89, 101113. https://doi.org/10.1016/j.wocn.2021.101113.

Clement, Tanya Elizabeth. 2024. *Dissonant records: close listening to literary archives*, Cambridge, Massachusetts: The MIT Press.

Didirková, Ivana, and Anne Catherine Simon. 2024. "Experimental Phonetics and Phonology". In *The Routledge Handbook of Experimental Linguistics*, edited by Sandrine Zufferey and Pascal Gygax, 21-37. Routledge.

Garellek, Marc, Matthew Gordon, James Kirby, Wai-Sum Lee, Alexis Michaud, Christine Mooshammer, Oliver Niebuhr, Daniel Recasens, Timo B. Roettger, Adrian Simpson, and Kristine Yu. 2020. "Letter to the editor: toward open data policies in phonetics: what we can gain and how we can avoid pitfalls". *Journal of Speech Sciences* (9): 3-16.

Garruccio, Alessandro. 2021. "Il caso del Boston College e del Belfast Project: 2001-2016. Le fonti orali citate in una controversia giuridica internazionale". In *Buone pratiche per la storia orale. Guida all'uso*, edited by Alessandro Casellato, 105-29, Florence: Editpress.

Ginouvès, Véronique, Francesca Biliotti, and Silvia Calamai. 2018. "Les archives sonores entre demande sociale et usages scientifiques: quelles modalités pour réutiliser les sources enregistrées?". In *La diffusione numérique des données en SHS. Guide des bonnes pratiques éthiques et juridiques*, edited by Véronique Ginouvès and Isabelle Gras, 169-94, Aix-en-Provence: PUF Presses Universitaires de Provence.

Gordon, Elizabeth, Lyle Campbell, Jennifer Hay, Margaret Maclagan, Andrea Sudbury, and Peter Trudgill. 2004. *New Zealand English: Its history and evolution*. Cambridge University Press.

Hay, Jennifer, and Paul Foulkes. 2016. "The evolution of medial /t/ over real and remembered time". *Language* 92(2): 298-330. https://doi.org/10.1353/lan.2016.0036.

Hutmacher, Fabian. 2019. "Why Is There So Much More Research on Vision Than on Any Other Sensory Modality?", *Frontiers in Psychology,* 10: 2246. https://doi.org/10.3389/fpsyg.2019.02246.

Nodari, Rosalba, and Silvia Calamai. 2021. "Degemination in marginal Tuscan speech: Temporal analysis in legacy speech data". In *Production and perception mechanisms of sound change*, edited by Daniel Recasens and Fernando Sánchez- Miret, 67-85. Lincom Europa.

Őhman, Carl. 2024 *The afterlife of data. What Happens to Your Information When You Die and Why You Should Care.* University of Chicago Press.

Parlangèli, Oronzo. 1972. *Scritti di dialettologia*. Galatina: Congedo.

Piccardi, Duccio, Fabio Ardolino and Silvia Calamai, eds. 2019. *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale / Audio archives at the crossroads*

of speech sciences, digital humanities and digital heritage. Studi AISV 6. https://www.studi.aisv.it/index.php/home/issue/view/12.

Piccardi, Duccio, and Silvia Calamai. Forthcoming. "Fear of FAIR? Towards a new Italian incentive to oral data curation".

Pievani, Telmo. 2018. *Homo sapiens e altre catastrofi. Per un'archeologia della globalizzazione*. Rome: Meltemi.

Prodi, Paolo. 2020. *Una storia della giustizia*. Bologna: Il Mulino.

Rogers, Irene, and Margaret Mcallister. 2014. "Ghosts in the archives: exploring the challenge of reusing memories". *Oral History in New Zealand*, 26: 11-19.

Scuola di Barbiana. 1967. *Lettera a una professoressa*. Florence: Libreria Editrice Fiorentina. (English translation: 1970. *Letter to a Teacher*. Translated by Nora Rossi and Tom Cole. London: Penguin Books).

Tafani, Laura. 2023. "La legislazione italiana: fonti, procedure, derive, effetti sul linguaggio". In *Il dovere costituzionale di farsi capire. A trent'anni dal Codice di Stile*, edited by Maria Emanuela Piemontese, 75-92. Rome: Carocci.

Tavolo permanente per le fonti orali. 2023. *Vademecum per il trattamento delle fonti orali*. Rome: Ministero della Cultura, Direzione generale Archivi.

Valentini, Cecilia, Duccio Piccardi, Silvia Calamai, and Maria Francesca Stamuli. 2023. "Da cornice a soggetto. Il documento sonoro nell'infrastruttura *Archivio Vi.Vo.*". *Archivi*, 18(1): 88-125.

Zagrebelsky, Gustavo. 2007. *La virtù del dubbio. Intervista su etica e diritto*, edited by Geminello Preterossi. Rome-Bari: Laterza.