



Citation: Piccardi, D. (2025). Less a census, more a lesson. A critical examination of the *Gra.fo Reloaded* oral archive inquiry results. *Oral Archives Journal*, 1: 133-172. doi: 10.36253/oar-3343

Received: October 15, 2024

Accepted: December 7, 2024

Published: February 20, 2025

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

ORCID

DP: 0000-0002-4985-4360

© 2025 Author(s). This is an open access, peer-reviewed article published by Firenze University Press and USiena PRESS (<https://www.fupress.com>) and distributed, except where otherwise noted, under the terms of the CC BY 4.0 License for content and CC0 1.0 Universal for metadata.

Landscaping Oral Archives

Less a census, more a lesson. A critical examination of the *Gra.fo Reloaded* oral archive inquiry results

DUCCIO PICCARDI

Università degli Studi di Urbino Carlo Bo, Italy
E-mail: duccio.piccardi@uniurb.it

Abstract. This paper recounts the oral archive Tuscan census promoted by the *Gra.fo Reloaded* project. With twenty-three responses over more than four hundred addressees, the census cannot be considered other than a failure whose productive potential is on hold. Through a complete report of the census aggregate data, this paper attempts to provide a glimpse of the current local situation, highlight present issues, and inform future actions. It also takes the opportunity to frame *Oral Archive Journal* as a place open to discussion regarding any aspect of failure in oral archive research.

Keywords: census, Tuscany, failure in science, action planning.

1. AN UNDESIRABLE INTRODUCTION

Given its intersubjective and experimental nature, Casellato and Lampe (2023) recently described oral history as a trial-and-error endeavour in which learning from mistakes constitutes a fundamental opportunity for disciplinary advancement. Of course, this implies a research community willing to disclose its missteps and put them on the table for public discussion. Curiously enough, providing an account of a conference held in this spirit, the authors noted that the topic of oral archives was by far the least apt to be tackled with this mindset, and hypothesised that potential contributors were held back by a feeling of embarrassment, which is supposedly specific to this line of research.

In actual fact, oral archives do not have a special place in the academic discussion on error and failure. As biologist Stuart Firestein (2016, 39–47) nonchalantly noted, the fact that failure is always around the corner in any scientific activity is guaranteed by science itself, i.e., by the Second Law of Thermodynamics as a

formal explanation of entropy. In recent times, scholars from disparate research areas have strived to normalise failure as a routinary component of scientific life (Parkes 2019), valuing failure as theoretically interesting (Barwich 2019), and envisioning outlets hosting academic discussion of failure (Verbuyst and Galazka 2023). Despite this trend, and possibly in relation to current practices of academic acknowledgement, researchers tend to lump together several undesirable aspects of their work, such as mistakes, errors, and even negative results and inconclusive findings, under the common label of stigmatised failure (Schickore 2021). This attitude has several drawbacks. Indeed, the inadequacy of policies and reward systems for the dissemination of errors has its social costs (Shur-Ofry 2016): given the centrality of tentativeness in the nature of science (Allchin 2012), errors should be systematically explored in order to become productive (Schickore 2005). Moreover, the stigmatisation of failure might be detrimental to the well-being of researchers, as the constant demand for excellence is one of the most widely reported themes in studies discussing mental health issues among those working in academia (Nicholls et al. 2022).

The current state of academic publishing reflects the issues described here. An influential study by Fanelli (2012) reported a robust increase in the number of papers containing positive findings between 1990 and 2007, to the detriment of negative results. While some studies comment on cases in which negative result papers are blocked off during manuscript review processes (e.g., Isbell et al. 2022), others suggest that this so-called publication bias acts primarily among authors, who are not willing to invest time and effort in the submission of works without positive outcomes (Franco, Malhotra, and Simonovits 2014; Van Lent, Overbeke, and Out 2014; see Dickersin 2005 for an overview of early data). Be that as it may, a recent linguistic analysis by Wen and Lei (2022) suggests that this stigmatisation of negativity is even substantially altering the way papers are written, with a constant increase in positive words used to communicate findings. Of course, various countermeasures are being implemented across several research fields, such as the constitution of outlets (Pfeffer and Holsen 2002) and recurrent workshops (Rogers, Sedoc, and Rumshisky 2020, III) specifically dedicated to negative results, or the implementation of preregistration or registered reports in the pipeline of academic publishing (e.g., Roettger 2021). Coming back to fields concerned with human voice, researchers in linguistics also stressed the importance of negative result disclosure: for example, Eddington (2008, 10–11) stated more than fifteen years ago that the acritical dismissal of negative results is a “common pitfall” in pseudoscience, while, more recently, Kortmann (2021, 1221–2) suggested that “the courage – both on the side of the authors and the journal editors – to publish ‘negative’ results” can save the field from developing symptoms of a quantitative crisis. While Kortmann’s overall perspective is one of cautious optimism, the wording of this passage might reveal that the fight against the stigmatisation of failure is still ongoing. Ideally, reporting something so intrinsically pervasive in any kind of human endeavour (Firestein 2016) should not entail any form of bravery.

What you have in your (digital) hands is the first volume of a journal stemming from various roots. As Silvia Calamai already narrated in her Editorial, some of them can be showcased with pride: indeed, OAr Journal results from years of fruitful collaborations between linguistics and oral history in the Italian Region of Tuscany, and from a cross-disciplinary work environment (the one leading to the *Vademecum per il trattamento delle fonti orali*, Tavolo permanente per le fonti orali 2023; henceforth *Vademecum*) open to the pondering of the solutions

for the academic acknowledgement of oral archive constitution (Piccardi and Calamai Forthcoming). Another one is in comparison far less desirable. Recently, we embarked on a new census of the oral archives produced in Tuscany. Despite our efforts to build the form following reliable pre-test procedures, and our dissemination plan including hundreds of addressees, the census spectacularly failed to reach its primary goal, i.e., to provide an updated picture of the documents conserved in Tuscany, by paradoxically gathering together fewer responses than previous regional inquiries. This result represented some sort of tipping point, which strongly led us to conceptualise OAr Journal as an attempt to incentivise the curation of oral documents and turn a deafening silence into a plural research community.

Mine is the task of leading you through this particular root of the journal, which is indeed a very dark one. The aim of this contribution is twofold. Firstly, I will look at the barely accessible sphere of Tuscan oral archives through the peephole offered by our respondents' data. Of course, the field of view will be very limited and any generalisation highly implausible; nonetheless, our census outputs may serve as the basis of future confirmatory inquiries to be conducted once the lock on the door starts to slacken. Secondly, I will try to make these less-than-desirable results productive by interpreting their patterns with the aim of advising forthcoming actions for the preservation of oral archives at the National level. This second aim will hopefully pave the way to frame OAr Journal as a place of discussion on the negative aspects of research and against their stigmatisation, whether caused by field-specific embarrassment (Casellato and Lampe 2023) or macroscopic changes in the academia as a whole. In the next paragraph, I will expound the census' underlying project (§2.1), the phases of its drafting, and our dissemination strategy (§2.2). We have already talked in detail about these latter issues in Piccardi and Calamai (2023, Forthcoming), so that the information reported here is expedient just to provide context, while adding original further reflections on specific aspects. The reader will find the whole census form in the Appendix. §3 will then delve into the actual results. Given the overall goals of this contribution, I decided (together with Silvia Calamai) to follow the presentation style of Cappelli and Rioda (2009), who gave priority to the analysis of individual census questions in order to provide aggregate estimates; this will also allow for a more focused discussion, which will be presented in §4. Lastly, conclusions will be drawn in §5.

2. THE *GRA.FO RELOADED* ESSENTIALS

2.1 *The project: reloading Gra.fo and walking sounds*

Gra.fo Reloaded is part of the research program on oral archives envisioned and promoted at the University of Siena. Run between June, 2022 and the end of November, 2023, the project was funded by Regione Toscana (*Fondo per lo Sviluppo e la Coesione – FCS 2014-2020, Giovanisì*), Siena University, Fondazione Sistema Toscana, and Ecomuseo della Montagna Pistoiese; moreover, it benefited from the participation of Soprintendenza Archivistica e Bibliografica della Toscana, Istituto di Linguistica Computazionale “Antonio Zampolli” (Pisa National Research Council) and CLARIN-IT. As its name implies, the main theme of *Gra.fo Reloaded* was one of revitalisation. Over the years, the access to the Tuscan oral archives digitised and processed during the original *Gra.fo* project (2009-2013; Calamai

2011) had become discontinuous at best because of infrastructural and managing issues¹. For this reason and following similar attempts to link the *Grafo* archives to more stable services (e.g., Calamai and Frontini 2016), *Grafo Reloaded* evaluated the feasibility of transferring its materials into a newer archival platform, *Archivio Vi.Vo.* The homonymous project (2019-2021) managed to finalise a restricted-access beta version including an archival pipeline leading the user from the description of a preservation master copy to the reconstruction of the individual oral documents (Calamai et al. 2022). Before blindly continuing with the development of the platform, *Archivio Vi.Vo.* was in clear need of beta testing: in this sense, *Grafo Reloaded* killed two birds with one stone.

In order to achieve some level of geographical coherence with the area of interest of the Ecomuseo della Montagna Pistoiese, *Grafo Reloaded* decided that the Anna Buonomini archive was the ideal test subject for this pilot transfer. The archive consists of four series of (twenty-one) documents gathered together in the *Provincia di Pistoia*. Two of them contain the original interviews pertaining to two research strands of the eponymous researcher, which were conducted towards the end of the 80s: the *Cappel d'Orlando* series, which is mainly about folktales and legends revolving around peculiar Appennine range sites (partially reproduced in Buonomini 1987); and the *Castagno* series, expounding local chestnut growing culture. The other two host more recent documents: a set of interviews of historical and ethnological interest conducted by high school students (*Studenti* series) and a lecture about the toponymy of the city of Pistoia and its surroundings (*Toponomastica* series). Archive reprocessing led to several improvements in the description of its contents, including time-aligned segments; upgraded metadata, especially with respect to the features of the (fourteen) analog carriers and the digitisation process; an updated archival terminology; and the (re)discovery of two additional documents. These belong to the *Studenti* and *Toponomastica* series and were presumably considered lost because of their low audibility², which was partially circumvented thanks to the expertise gained during the *Archivio Vi.Vo.* project. With respect to the platform beta testing, this operation pinpointed several areas for improvement: in particular, it was made evident that the implementation of metadata schemes for certain archival levels was still too embryonic. Thus, the completion of the archival structure of the platform was set as a priority for any of its future developments.

In direct connection with this line of action, *Grafo Reloaded* teamed up with the Ecomuseo della Montagna Pistoiese in order to conceive territorial promotion strategies based on the reuse of the Anna Buonomini documents. In actual fact, Pozzebon, Biliotti, and Calamai (2016) had already argued that the *Cappel d'Orlando* series of the Buonomini archive could be suitable for an augmented cultural heritage mobile application that reproduces audio files when the user physically approaches a list of sites in the Pistoiese Appennine range. This concept was further refined by referring to the soundwalk research tradition (e.g., Bradley 2016): in particular, *Grafo Reloaded* explored the possibility of building a series of listening

¹ It should be noted that, at the time of writing this paper, the *Grafo* original portal is up and running again, waiting for the transfer operation piloted by *Grafo Reloaded* (see below) to be completed (<https://grafo.sns.it/site/about>, accessed July 1, 2024).

² See the Buonomini archive description in the original *Grafo* portal, which states that the *Toponomastica* document is “non molto intellegibile” (<https://grafo.sns.it/site/archivi?ArchivioSearch%5Bnome%5D=buonomini&ArchivioSearch%5Bdescrizione%5D=>, accessed July 1, 2024).

spots disseminated over the path leading from the small village of Pian degli Ontani to the sandstone pinnacle named Torre del Fattucchio, centred around the local legend of a treasure hidden in its area (Lipparini 1930). Both a field inspection and preparatory work to select a platform suitable for hosting the Buonomini documents were conducted. The latter resulted more puzzling than we had previously thought, since the most refined soundwalk services were either non-compliant with the requirements of the European General Data Protection Regulation or hardly sustainable in the long run, due to specific business models. After several consultations with the University of Siena Data Protection Officer, we ended up drafting a data processor appointment agreement document, which was stored as a model for any future similar endeavours. However, due to the premature end of the project, the actual Buonomini soundwalk has yet to see the light of day.

2.2. *The elephant in the room: the Grafo Reloaded census*

The design, testing, and dissemination of a new census of the oral archives produced in the Region of Tuscany occupied the bulk of the time spent on the project. As a matter of fact, the idea of conducting such an inquiry was not by any means new. After the seminal Andreini and Clemente (2007), both Cappelli and Rioda (2009) and the original *Grafo* project (Calamai 2011) ran regional censuses. Moreover, with different means and scopes, at least three other lines of work investigated the distribution of oral archives at National level: Barrera, Martini, and Mulè (1993), Benedetti (2002), and, *mutatis mutandis*, Galatà and Calamai (2019). Other than for the sake of updating the existing data, *Grafo Reloaded* tried to carve its own niche by grounding itself on the archival concept refined during the *Archivio Vi.Vo.* project (Valentini et al. 2023) and being originally receptive to digital-native archives. Moreover, it tried to answer to one of the issues brought up by other contemporary census efforts (Michetti 2022), i.e., how to lower the costs of periodically refreshing our knowledge about the production of oral archives and the rate of obsolescence of this kind of information, by striving to build a research community willing to invest time in the curation of oral materials. To this end, the *Grafo Reloaded* census made its respondents aware of a set of dedicated tools and initiatives, which were either already available or in preparation: the *Vademecum*, the *Archivio Vi.Vo.* platform, and what was, at the time, a first vague sketch of the journal that you are reading right now. With respect to the *Vademecum* and *Archivio Vi.Vo.*, participants were able to join the respective mailing lists through the census form in order to keep themselves up to date on the latest developments and events about oral archives in Italy.

The first step in our census form drafting procedure consisted in the critical selection and adaptation of elements taken from all the lines of oral archive inquiry mentioned above. Thanks to the experience gained during the *Archivio Vi.Vo.* project, we put particular emphasis on aspects related to any eventual restoration of physical carriers, the digitisation of their contents, the legal issues surrounding the oral archive, and the presence of other kinds of allied documents in it. In this phase, through the network centred around the *Vademecum*, we managed to discuss the draft with another Italian oral archive census endeavour (Michetti 2022). Following this, we dived deep into the survey design literature in order to decide on suitable solutions for data collection modes and question structuring. We ended up choosing a web-based strategy and implemented the draft into the intuitive and widespread Google Forms. For

the sake of easing the burden of completing the census, we made use of the platform's skip logic functionality; moreover, we valued question structures that enhance respondents' perceived usability of the form; lastly, we added the built-in progress indicator for transparency.

The Google Form draft of the census was pre-tested following the guidelines of Statistics New Zealand (2018) and the United States Census Bureau (Nichols et al. 2020). Thirteen subjects of very diverse levels of education and involvement with oral documents partook in a qualitative cognitive interview and were asked to fill in the census form while thinking aloud in front of the author of this paper. During this step, notes were taken about their interactions with the form and comments on perceived inconsistencies, ambiguities, missing options, etc. Subsequently, participants completed a shortened version of the Questionnaire of User Interface Satisfaction (QUIS; see Rao and Mingay 2001 for the version adapted in Italian for this pre-test), consisting of eleven five-point Likert items on specific aspects of filling in the census followed by three items on general impressions. This procedure led to forty-four distinct observations on the census form mostly pertaining to formatting issues, both formal and content revisions to specific questions (or options), and clarification requests. The prevalence (31) of the latter was coherently reflected by the QUIS results: the only mean score which was under 4 out of 5 pertained to the item on the census appropriateness to inexperienced users. Overall, the pre-test signaled to us a clear call for simplification. With this aim, we built a revised, more user-friendly census form featuring a slight increase in the number of pages (49, up from 47) and questions (100, up from 98). It should be noted that our skip logic implementation (most of the time) drastically reduced the effective number of items encountered in each specific completion run³, and perceived completion speed was deemed acceptable by our pre-testers (4.3 out of 5). The final census form presented the following main sections, with the respective total number of items between brackets: physical carriers (11); digital documents (13); other archive features (archival arrangement, description, long-term preservation; 8); document contents (12); other related elements (allied documents and transcriptions/annotations; 7); legal issues (9); publications (2); prior knowledge of pertinent topics (6); declarations of interest (7); personal information (on all the subjects involved, including information about archive location and accessibility; 25). These sections were prefaced with a brief introduction regarding the contents and the scope of the census (including a link to a more extensive document) and a disclaimer concerning personal data treatment (including a declaration of consent option) redacted by the University of Siena legal office, in this order.

We then agreed upon two consecutive phases of census dissemination. The first relied on a circular cover letter sent to the offices of all the Municipalities of the Region of Tuscany, a selection of cultural foundations and Italian scientific associations dealing with human voice, the coordinators of Tuscan Ph.D. programs dedicated to pertinent topics, and a hand-picked list of individual researchers pinpointed through a scrupulous examination of the publications of all the Tuscan academic personnel filtered by a set of related academic fields. The link to the census form was posted on the websites of all the involved partners, disseminated through in-person events and even had some media coverage. After this first phase, we contacted other potential respondents through personalised emails, phone calls, and social media accounts. These also included U.S. Italian cultural foundations and ex Master of Arts or Ph.D.

³ A minimal run consists of approximately sixty questions.

students selected through an inspection of the thesis repositories of the Tuscan Universities. Upon request, in-person meetings with both the personnel of foundations and individual researchers were also held with mixed results. In total, we reached around 420 addresses – even though having an exact estimate of the census reach is actually impossible, since we do not have data on the effectiveness of the dissemination strategies of all our partners and we cannot ascertain whether unanswered emails have actually been read or not (because of, e.g., obsolete addresses or unsignalled technical issues). The census form officially accepted answers between December 10th, 2022 and November 30th, 2023. At the end of the project, we received 23 oral archive descriptions⁴.

Even taking into account known difficulties in gathering answers in similar circumstances (e.g., Ferriter 1998, 91), this total tally seems underwhelmingly low and disproportionate to the numbers of our dissemination phases. As I mentioned above (§1), all previous Tuscan oral archive censuses managed to end their runs with more responses than we got: specifically, Andreini and Clemente (2007) contained more than 120 oral archive profiles; Cappelli and Rioda (2009) collected 41 responses out of 107 contacted addresses; the original *Grafo* project (Calamai 2011) described more than 30 archives. Of course, these past inquiries are not easily comparable to the *Grafo Reloaded* census in terms of overall available resources. In particular, both Andreini and Clemente (2007) and the original *Grafo* were able to systematically arrange in-person meetings and interviews with their respondents, which may be considered the optimal, albeit expensive, solution to achieve high response rates (Vannette and Krosnick 2018, XXI). Conversely, Cappelli and Rioda (2009) adopted a more indirect approach, and arranged meetings only upon specific requests. However, they focused on a previously verified (Benedetti 2002) set of cultural institutions, while the *Grafo Reloaded* census was largely of exploratory nature.

Other formal aspects of our census could have led it to its demise, despite the pre-test phase described above and the multiple occasions for its peer revision⁵. The extensiveness of the census and the consequent high completion time are the prime suspects. Intuitively enough, several studies reported a negative correlation between stated/actual survey length and start/completion rate (e.g., see the reviews in Liu and Wronski 2018; Sammut, Griscti, and Norman 2021, 19). Moreover, after asking a large number of web respondents for their opinions about maximum survey length, a recent work ended up reinforcing informal guidelines and suggesting a “20 minutes rule” (Revilla and Ochoa 2017). During our pre-test, we observed actual completion times ranging from 30 to 60 minutes. The positive responses to the QUIS and the relative absence of length among the topics discussed with interested parties both before and after

⁴ Note that, initially, we envisioned a series of in-depth interviews to researchers at very different stages of their careers to provide a qualitative counterpart to the census. Two of these were performed (see the Acknowledgements section), but it soon became evident that our original intents were not adequately backed by the actual public response to the census. In other words, we could not link these individual stories to generational (or disciplinary) patterns if the census was not able to provide them in the first place. We hope that we will be able to repurpose these rich narratives in future endeavours.

⁵ The census form was publicly discussed before and after its publication through two oral communications coauthored by Duccio Piccardi and Silvia Calamai: “Il Progetto *Grafo Reloaded*”, presented at the conference *Un anno di ‘Vademecum per il trattamento delle fonti orali’*, Rome, October 27th, 2022; and “How many archives are there in your home? Piloting a new Tuscan census in the *Grafo Reloaded* project”, presented at the XII AIUCD conference, Siena, June 6th, 2023, which was later published in Piccardi and Calamai (2023) through a peer review process.

the census launch probably made us underestimate this aspect. In addition to that, given the ambiguous effect of progress indicators on survey drop-offs (Villar, Callegaro, and Yang 2013), the presence of this design element in our specific census may have backfired. Again, direct comparisons with other similar Tuscan censuses are hardly feasible. Indeed, Cappelli and Rioda (2009) grouped their questions into a total of 35 items, which is far less than ours. Andreini and Clemente (2007) relied on four forms per archive, the length of which varied depending on the number of analogic carriers. In their case, the data collector, and not the respondent, filled in the forms; nonetheless, one specific form about the archive history required the respondent to partake in an in-person interview with the data collector, which often included autobiographical elements and was thus quite lengthy (Simonetti 2007, 275-6). As mentioned above, the original *Gra.fo* inherited this *modus operandi* and developed a 7-question interview structure named “Say something about your archive” submitted to the owners of archives with little to no attached documentation (Calamai and Biliotti 2017a, 87-8). Luckily enough, these can be streamed through the *Gra.fo* portal, and lasts up to approximately 45 minutes, which falls in the middle of our completion time range. However, differences in task modality (such as, in our instance, online and self-administered vs. in-person and dialogic) may presumably trigger changes in the factors affecting subjective time perception, such as engagement and attention (see Thönes and Stocker 2019 for a review). Moreover, the presence of an in-person interviewer interacts with social desirability and satisficing dynamics and may ultimately lower the number of quick and low-effort responses given to a survey (e.g., Atkeson and Adams 2018), naturally increasing the time needed to complete interview inquiries. Lastly, as a recent Italian study suggests (Grimaccia et al. 2023), people preferences for web- or interview-based censuses are tied to several socio-demographic variables, so that different subsets of the population may presumably show more or less tolerance towards extensive and time-consuming forms depending on their proposed mode of completion.

As the *Gra.fo Reloaded* census form is freely available, we encourage future iterations to develop a shortened version, and/or tweak other formal aspects, in order to pinpoint its major weaknesses and make the best of this work. In the meantime, we strived to interpret the fragmentary glimpses provided by personal communications and demographic patterns found in the pool of actual respondents (see §3.8, 3.10) in order to understand whether the difficulties surrounding the census may have also depended on specific reactions on the part of the addressees. We identified a triad of major underlining themes, which are also recurrent in the literature focusing on the reasons for open science noncompliance (Gomes et al. 2022): knowledge barriers, lack of incentives, and trust issues. The first concerns deep uncertainty about archival and legal aspects, which ultimately dissuaded the respondents from even trying to open the form. In other words, respondents told us that they were unsure of their entitlement to fill in the census, because of doubts on ownership or content disclosure. The second is somehow related to completion time: the few responses received by full-time active permanent academic personnel, together with their comments, led us to hypothesise that oral archive curation is understandably unattractive to those who are already burdened by a heavy load of teaching, research, and administrative duties. Lastly, the third was expressed through fear of the negative consequences of disclosing potential flaws (see above, §1), and concerns about being induced to transfer rights, make data public, or deposit the archive in specific repositories. In actual fact, this latter theme is deeply intertwined with the

issue of trust in contemporary society (Hardin 2006) and the progressive endorsement of open science (Stracke 2020) but goes well beyond the field of action of a regional project on oral archives. By contrast, we argued that the recently published *Vademecum* could be seen as an asset to lower the knowledge barriers of oral archive curators in Italy and the relative probability of their being stuck in decision paralysis. Finally, in our own small way, we hope that OAr Journal, and, specifically, its “Oral data curation” section (see the Editorial), will be of help in providing researchers with suitable incentives to take care of their oral archives. In fact, by accepting overlay papers in which oral archives are reviewed and published together with them, OAr Journal follows in the footsteps of other initiatives finding a place for this type of effort in the academic acknowledgement system (e.g., Fitzgerald 2021).

From its inception to its future legacy, this section provided a bird’s-eye view on the *Grafo Reloaded* project, its census, and its general shortcomings. In the next section, I will report on the actual census results. The original item order will be followed, and aggregate data will be presented; qualitative observations and discussions on their relevance to the understanding of the situation of oral archives in the region of Tuscany, the census dynamics, and effective strategies for the planning of future actions and lines of research can be found in §4.

3. THE *GRA.FO RELOADED* CENSUS RESULTS

This section contains a detailed report of the results of the *Grafo Reloaded* census. Before diving into things, a few remarks are in order. Firstly, respondents to previous Tuscan censuses were allowed to partake in our survey: therefore, part of the materials discussed here may already have been considered in previous similar endeavours. Indeed, we noted that two of our respondents were involved in Andreini and Clemente (2007), another one in the original *Grafo* census and a fourth one was involved in both. As will be shown below (§3.1, 3.2), respondents may consider their archives open to new additions, so that having new contributions from these researchers does not necessarily end there. Moreover, our census did not explicitly target institutional archives; nonetheless, some of our respondents ended up describing archives linked to public or private institutions, associations, activities, projects, et sim. These were the Accademia della Crusca; laboratory/archive Alberto M. Cirese – LIDE-AV-Ars Videndi, University of Siena; Associazione Aureo Anello; Centro Studi Tradizioni Popolari Toscane; Film Documentari d’Arte; Istituto italiano Festival dei Popoli; the branches of Florence and Prato of the Italian General Confederation of Labor – CGIL; and the project “Officina delle voci. Atelier della creazione giovanile”. Some of these archives are extraordinarily larger than those of individual researchers and consequently, they will be explicitly nominated whenever their numbers substantially skew the aggregate picture. Lastly, note that our twenty-three archives were described by twenty-two respondents: for this reason, items related to the respondents will add up to the latter total number, especially in §3.9 and 3.10.

3.1. *Physical carriers and restoration*

The first check-all-that-apply (henceforth, CATA) item investigated the type of physical carriers contained in the archives. The *Grafo Reloaded* census adopted a very broad accepta-

tion of this category of objects, which pertains to all the physical means of local storage of the documents, from cylinders to solid-state drives. All the proposed options were flagged at least once, with the sole exceptions of “Other disc records” and “SD cards”. Apparently, the most common oral document carrier type is the compact cassette (10 out of 23 archives), followed by external (8) and internal (7) hard disk, DVD (6), Audio CD (5), CD-ROM, USB flash drive, and 1/4 inch tape (4 each), 1 inch tape (3), post-1925 78 RPM, 45 RPM, and 33 RPM discs, 1/2 and two inch tape, other magnetic carriers⁶, and DAT (2 each). The other options (including cylinder and magnetic wire) were selected once; moreover, three archives contain other types of video recording carriers (such as VHS cassette).

The second item required participants to quantify their physical carriers. One researcher did not provide details on this aspect, the total number of physical carriers equalling to approximately⁷ 4,417 units. Of course, the number of carriers was greatly variable among individual archives. In particular, ten archives consisted of digital documents only, thus presenting one to six carriers (i.e., digital data storages). Conversely, the archive containing the audio-visual documents of the Florentine branch of the Italian General Confederation of Labor (CGIL) presented the greatest number of carriers (1,874), followed by the archive of the ethno-anthropological branch of the Department of Philosophy and Social Sciences of the University of Siena (1,455). All other non-digital-only archives (10) were much smaller in this respect, ranging from 3 to 386 carriers. The third item expanded on this topic by asking participants to estimate the number of carriers per type. Since this topic was dealt with at an even greater level of numerical approximation, I will group some types and rank them by their estimated quantities. Coherently with the ranking of item 1, compact cassettes were in first place (more than 2,000); these were followed by video-only carriers (e.g., VHS, MiniDV, etc., DVDs not included: more than 1,200), CDs (both audio and ROM) and DVDs (more than 750), tapes of various types (more than 200) and discs (more than 100), external and internal hard disks (more than 30), and less represented types of carriers. With respect to the fourth item, the majority of the archives (15) were not open to other physical carriers, while three forms took the opposite stance, and seven did not provide an answer.

The fifth item explored the state of conservation of the carriers through a four-level scale. With respect to the carriers, none of the archives reported poor condition; conversely, most of carriers were in near mint condition (12), six in very good, and five in good condition. Note that this taxonomy was adapted from Popoff (2009) and, despite the positive valence of its labels, implies the presence of issues in the reproduction of audio documents, which were explicated on the form in brackets. For example, very good carriers show minor wear with little impact on the audio quality, while good carriers manifest wear which is noticeable and detrimental to the quality of the reproduction. By not selecting the “poor” label, participants told us that their carriers were not heavily damaged, and that eventual wear did not hinder the reproducibility of the documents. The following item (number 6) asked participants if this overall label was suitable for all the carriers or if specific subsets were in significantly different states of conservation. This follow-up let us know that, in actual fact, two post-1925

⁶ Note that, at least in one of these two cases, this option was selected to express the presence of tapes without any further specifications.

⁷ Many respondents gave approximate estimates instead of precise numbers; when a numeric range was provided, a rounded down average was computed.

78 RPM discs were in poor condition; and that the state of conservation of the various tapes contained in an institutional archive was assessed through visual inspection alone, since the institution did not possess any play-back device. A similar statement was made with respect to a private research archive in the following block of items (7-11) on physical carrier restoration. Only one respondent declared relying on restoration procedures: in particular, the researcher stated that the compact cassettes in question underwent some unspecified tape cleaning process of which there was no documentation available; moreover, it was contextually noted that the respondent no longer had any device to listen to the contents.

3.2. *Digital files and digitisation*

The twelfth item moved the focus of the census from the physical to the digital realm: leaving aside the analog carriers, do the Tuscan archives also contain digital files? Most of them (21) indeed do. Item thirteen attempted to provide a rough quantification of these; however, seven of the twenty-one forms did not contain any answer regarding this point. The approximate total number of digital files contained in the remaining fourteen archives is 658,290 units but again, as for the physical carriers, their distribution is vastly uneven. A single private research archive counted alone approximately 650,000 units pertaining to years of video documentary efforts; the digital-native archives of two young linguists contained 3,092 and 2,640 units, respectively; the already mentioned large archive of the CGIL Florentine branch hosted approximately 2,000 units which, however, admittedly included also documents from the Pistoia branch. The other archives were, in this respect, far smaller, with numbers ranging from just 3 to 287 units. The following item (14) investigated the digital formats of these files. The most common digital format type contained in the Tuscan archives is .wav (15 out of 21 archives), followed by .mp3 (9), .mp4 (3), .avi (2), .ogg, .mod, .mpeg, and .mts (one single occurrence each). The fifteenth item was a digital counterpart to item 3 and asked for a quantification of the files per format type. Unfortunately, respondents did not provide any answers to this question with respect to ten archives, so that the remaining information is not easily interpretable. Probably the most interesting pattern related to this item is that, whenever both .wav and .mp3 were mentioned in the same answer, .mp3 files were considerably fewer than the ones in .wav (90 vs. 10, 3,088 vs. 4). Similar to item 4, item 16 revealed that the majority of the archives (11) were not open to other digital files, while five forms declared the opposite, and the same number did not reply to the item.

Glossing over item 17 about audio quality (bit depth/sampling rate), which was answered by only six forms, our respondents were split on whether their digital files were the outcome of a digitisation process or not (item 18, 11 yes; 10 no). In four cases, these digitisation processes were started during the Nineties, while just in one single case was the process started very recently in the 2020s; with respect to two archives, respondents were not able to provide information on the year(s) of digitisation (item 19). Most of the archives (item 20; 8 out of 11) contained audio digitised by the respondents themselves, while in three cases the digitisation process was entrusted to third parties. Only two of these procedures left written documentation (item 21), and only in one institutional case was this kept together with the archival materials (item 22).

Lastly, respondents were asked if their files were also stored online (item 23). Affirmative answers were gathered only in six out of twenty-one forms. Curiously enough, with the sole exception of the two CGIL archives relying on the same servers of the Confederation, respondents indicated different solutions, including YouTube, Dropbox, OneDrive, and other websites (item 24).

3.3. *Archive arrangement, description, long-term preservation, and usage*

The following batch of items dealt with aspects of archive curation and actual use. Firstly (item 25), most of the archives underwent some form of arrangement (14 out of 23). Respondents arranged their materials by subject, author, place of conservation (4 each out of 14), or following numerical (3), alphabetical, chronological criteria (2 each), or even by the interviewee's place of birth and sex, and the type of interview, the type of carrier, phonetic/phonological, and other idiosyncratic criteria (1 each; item 26). Subsequently (item 27), respondents were asked if their archives were described: most of them indeed were (16 out of 23). However, the method used to describe (item 28) half of the archives (8) was not disclosed; a variety of strategies were deployed for the remaining archives, including the indication of the interviewee and interviewer names, place, date, and duration of the recording (and other similar solutions); recording context, method of interview, project name and content synopsis uploaded as YouTube descriptions; contextual information integrated in the transcription of the contents; information reported on the physical carriers. Almost none of the archives could rely on planned actions for the long-term preservation of the carriers and the digital documents (item 29); only in one single case (out of 23) did an independent researcher tell us that he periodically backed up his files in several digital data storage devices (item 30). The last items in this section investigated how often the archive materials were listened to (item 31) and when was the last time that someone had listened to them (item 32). With respect to the first question, the majority of the archives were apparently never (or almost never) listened to (12), or only one or more times per year (8), while three archives were listened to at least once a week, a month or every six months (1 each). Lastly, eleven archives were last listened to more than one year before the census, four between six months and one year before, three between one and six months before, and between one week and one month before, and two were listened to in the same week as the census.

3.4. *Say something about your archive (contents)*

The next subsection was named after the original *Gra.fo* interview protocol (see above, §2.2) and contained items on the context of the recordings and their contents. Firstly, (item 33) participants were asked to provide details on when the recordings were done. Three archives contained recordings dating back to the Fifties, three to the Seventies, two to the Eighties, four to the Nineties, four to the first decade of the new millennium, four to its second decade, while the remaining three archives showcased very recent recordings only (2020 to present day). From recording time to recording place, item 34 investigated where the recordings were made. Unfortunately, we do not have this information with respect to two

archives. The remaining twenty-one archives contained recordings made in all the Tuscan Provinces. Here is a list of the places pinpointed by our respondents, in alphabetical order:

- Arezzo (3): Arezzo, Bibbiena, Carda, Castello di Sorci, Castel San Niccolò, Chitignano, Levane, Monterchi, Poppi, Raggiolo, Rassina, Terranuova Bracciolini.
- Florence (9): Castelfiorentino, Florence, Impruneta, Marradi, Palazzuolo sul Senio.
- Grosseto (3): Grosseto, Massa Marittima, Orbetello, Ribolla.
- Leghorn (2): Campiglia Marittima, Leghorn, Piombino.
- Lucca (3): Lucca, Viareggio.
- Massa-Carrara (3): Bedizzano, Bergiola, Caprio, Carrara, Colonnata, Filattiera, Filetto, Gigliana, Groppo, Iera, Lusignana, Mulazzo, Nezzana, Pieve, Ponticello, Pontremoli, Rocca Sigillina, Tresana, Treschietto, Villafranca in Lunigiana.
- Pisa (4): Buti, Pisa.
- Pistoia (6): Abetone, Agliana, Castelvecchio, Pescia, Pistoia, Ponte Buggianese, San Quirico.
- Prato (3): Montemurlo, Prato, Vaiano.
- Siena (5): Abbazia San Salvatore, Chianciano Terme, Gaiole in Chianti, Siena.

Other places outside Tuscany were also reported: Milan, Naples, various places on the La Spezia side of Lunigiana, Bari, Modena, Desenzano del Garda, Rome, Genoa, Bologna, and other more vague indications (Sicily, Sardinia, Trentino, African and South American countries).

Item 35 asked participants to state the disciplinary field(s) to which their archives were pertinent. Oral history was selected nine times (out of 23), followed by phonetics and phonology (8); dialectology, sociology (6 each); sociolinguistics (5); anthropology, musicology (4 each); political science (3); literary and poetry studies, archaeology, cognitive linguistics, art history, and documentary film studies (1 each). After this, item 36 requested the participants to list all the speech varieties spoken in the archive. As far as local varieties and dialects are concerned, the reader can get an idea of their geographical distribution by looking at the responses to item 34⁸. Here I will group these answers by their mention to standard Italian (7 out of 23), local and supra-local varieties of Italian (5), dialects (9), and non-Italian languages (3; these include English, French, Chinese, and Romanes).

The next item (37) dealt with the spoken text types contained in the archive. The most common type was life story (13 out of 23), followed by account of events, interview (12 each); spontaneous conversation (7); readings, conferences, *contrasto in ottava* (6 each); linguistic questionnaire, poetry (5 each); lullaby, political discourse (4 each); semi-spontaneous dialogue, narrative song, begging song, nursery rhyme, proverb, fable/novella, *maggio drammatico/epico* (3 each); ethnographic questionnaire, picture/object description, cumulative song (and similar types), lyric-monostrophic song, burlesque poem, riddle, invocation and exorcism, theatre, folk theatre, performative events (2 each); folkloric questionnaire, ethno-text, *zingaresca*, tongue-twister, legend, traditional medicine, ritual, *bruscello*, *vecchia/segala-vecchia*, Carnival-related events, concert, migration song, book presentation, inauguration speech (these last four types were found added under the “other” option; 1 each). This noticeable variety of types relates to heterogeneity of the themes emerging from the contents of the recordings (item 38). The list of potential options in this CATA was extraordinar-

⁸ Two notable exceptions are recordings in Bagitto (the variety of the Jewish community of Leghorn) and in some dialects of Corsica (spoken by seasonal workers).

ily extensive (see the Appendix), so that only the options selected more than twice will be reported here. Suffice to say that six respondents felt the need to add other options to the form in order to provide a more precise representation of the contents of their archive: this was especially true when the recordings were centred around the reading of specific excerpts selected for experimental purposes. With respect to the already available options, the archives contained materials with the following themes: work, the environment (9 each out of 23); anecdotes (8); everyday life, food and typical dishes (7 each); songs and folk music, conferences and debates, agricultural work, community life, biographies (6 each); fascism and antifascism, emigration, craftsmanship, the home and household activities, political activism, agriculture (5 each); farming culture, dialects and language varieties, work in the woods, dialect lexicon, womanhood, post-war period (Second World War), labouring work and life, professions, traditional crafts, traditions, celebrations, clothing (4 each); political meetings, religion, childhood memories, interculturalism, political history, immigration, Carnival, commemorations/anniversaries/ceremonies, Maremma, Nazism and racism (massacres, deportations), Resistance, sharecropping family, rural house, ages of life and social relations, animals, radio or tv recordings, teaching activities, demonstrations (3 each).

The following items consisted of two open questions about how and why the archive came to be (number 39) and to what aims (number 40). These included idiosyncratic answers which are hard to summarise here. Among the many specific research goals and institutional missions of analysis and preservation, we noted that in one single case more intimate reasons, like making family memories, were also reported. Item 41 then delved into the research methodology of the respondents asking them to talk about their interviewee recruitment strategies. One participant did not reply to this point; of the remaining twenty-two archives, we can highlight the following trends: three archives presented multiple recruitment methods; the same number of forms implied that the interviewer had no previous connections with the local community, while the opposite was true for eight forms; four archives contained recordings made thanks to the role of institutions in the recruitment process, these either being directly involved in the project or acting just as an intermediary; the same number of forms mentioned snowball and word-to-mouth sampling procedures; lastly, in three cases the respondent mentioned some kind of public announcement, either via social networks or flyers. In the same vein, item 42 investigated the setting of the recordings contained in the archives. Similarly to the previous item, one form did not contain any details regarding these aspects; twelve (out of 22) mentioned private houses; six listed various indoor spaces (schools, libraries, pubs, retirement homes, conference halls, etc.), leaving aside laboratories, which made an appearance in two forms; four listed outdoor spaces (e.g., squares); and the same number made reference to events (conferences, poetry sessions, etc.).

The last topic addressed by this section concerned funding issues. Only six out of twenty-three archives contained recordings which were made possible thanks to some form of funding (item 43); apparently, all these projects were mainly funded by public agencies (item 44).

3.5. Other related documents and transcriptions

The fifth section of the census focused on archive documents which were not (exclusively) audio but were allied to the oral documents described in the previous sections. Moreover,

it investigated transcription practices. Item 45 searched for non-audio-only allied documents in the surveyed archives and found them in thirteen out of twenty-three forms. Ten forms (out of 13) contained allied images or photographs, eight videos, and three textual documents (item 46). Most of the respondents did not attempt to precisely quantify the number of these allied documents (item 47): only in three cases was the total number of videos estimated, this being 67, 30, and around 10-15. Curiously enough, this item served as a sneak peek into the sub-types of non-audio documents, which included the actual books presented during the recorded presentations, photos for the sake of video-editing, as well as private ones, and related video interviews. Moreover, most of the archives (12 out of 23) contained documents which were transcribed and/or annotated (item 48). Item 49 revealed that this processing was, in the majority of cases, performed by the respondents themselves, although we also gathered sporadic references to co-authors, university students, and other related individuals. Of course, not every document in the archive was necessarily transcribed and/or annotated, this being the case for only two archives (out of 12). Four forms stated that approximately 80% of the archive documents were processed; three indicated that only half or less of the documents were annotated and/or transcribed; and the same number did not provide details on this point (item 50). These annotations/transcriptions were in hardcopy format, or stored in .doc (7 each), .TextGrid (PRAAT software) (4), .txt, .pdf, and .eaf (ELAN software) files. One participant even specified that the text files were annotated using the CHAT model of MacWhinney (2000) (1 each; item 51).

3.6. *Legal issues*

The sixth section of the census was dedicated to the legal aspects of the archives, from their formal recognitions of cultural value to potential issues related to the contents of the documents. Item 52 asked participants if their archive was declared of “particularly important historical interest” following Lgs. D. 42/2004 (see e.g., Visconti 2019, 169). Three respondents answered affirmatively; however, the following item (53) made it clear to us that this question had been freely interpreted⁹. After this (item 54), the census investigated potential copyright issues concerning the archive documents. Fifteen forms did not contain any clear answer to this point, while only three of them stated that the materials were without doubt free of any copyright issues. The remaining five were asked to provide some indications on who held the rights (item 55). Unfortunately, one respondent opted out of this question, while another possibly misinterpreted it; the remaining three forms indicated television and cinema producers as the rights holders of published documents, or the respondent himself. Item 56 moved on to the presence of personal data in the archive documents: seventeen archives did contain it. The same was investigated with respect to the presence of special categories of personal data (Art. 9 GDPR; item 57): this item revealed less clear-cut answers, with four “not available” options, nine “yes” and ten “no”. Only eight archives contained recordings made through the signing of an informed consent form (item 58), these being kept together with the archive in three cases only (out of 8; item 59). These eight respondents were also redirected to item 60,

⁹ Indeed, one respondent used this item to state that his archive appeared in the census by Andreini and Clemente (2007), while another specified the date of its indexing in the *Sistema Documentario Integrato dell'Area Fiorentina*.

where they were asked to pinpoint the legal aspects regulated through the consent forms. The most recurring aspects (7 out of 8 each) were goals of personal data processing and voluntary participation, followed by privacy, data processing techniques, data processing responsibilities (6 each); eventual reuse (3); compensation, complaint process (2 each); intellectual property and copyright/performer's rights (1 each).

3.7. *Publications*

This small section dealt with the forms of publication involving the Tuscan oral archives. The contents of the majority of the archives (18 out of 23) were indeed the object of publication (item 61). The most common type was scientific paper, which was reported in eleven forms (out of 18), followed by monograph (6); newspaper article, web post (including YouTube videos), catalogue, research dossier/report (2 each); magazine article, vocabulary entry, radio program, documentary film, museum set-up, and published song (1 each; these were added by the respondents under the "other" option; unpublished theses or dissertations were not considered here; item 62).

3.8. *Prior knowledge*

The remaining sections of the census were centred around the people involved with the archives. In particular, this eighth part asked questions about prior knowledge concerning the technical, archival, and legal issues of whoever actually made the recordings. Item 63 was focused on prior knowledge regarding technical aspects, such as recording tools, audio formats, etc. Three forms did not contain answers on this point, while thirteen expressed affirmative answers (vs. 7 "no"). These thirteen forms also contained some details on how this knowledge was acquired (item 64): seven reported university courses and six individual study; private courses, summer schools, workshops on oral documents, vocational training courses, and teaching by a thesis supervisor were also mentioned one time each. Item 65 moved on to prior archival knowledge. Leaving aside three "not available" responses, only two forms answered affirmatively (vs. 18 "no"). In this respect, one form highlighted the role of university courses, while the other of workshops on oral documents (item 66). Lastly, only one individual had prior legal knowledge (vs. 5 "not available" and 17 negative answers; item 67), which was acquired through workshops specifically dedicated to oral documents (item 68).

3.9. *Declarations of interest*

As was briefly mentioned in §2.2, one of the goals of the *Grafo Reloaded* census was to inform the research community about a series of tools and initiatives concerning oral archives. Accordingly, this section was specifically dedicated to this aim. Item 69 asked participants if they were aware of the existence of the *Vademecum*, to which we received only four affirmative answers (out of 22 individual respondents). Consequently (item 70), we asked participants if they would be interested in remaining up to date on the initiatives promoted by the *Vademecum* coordination group and join its mailing list, to which sixteen

replied affirmatively. The next three items tried to spread awareness about the *Archivio ViVo* platform and provide suggestions for its future developments. Firstly (71), participants were asked if they were generally interested in depositing their (digital) archives in online repositories. Fourteen of them replied that they were. We then provided a CATA list of features of online repositories asking all participants to indicate the ones that would incentivise them to deposit their archives (72). An intuitive graphic user interface was the most commonly selected feature (12 out of 22 respondents), followed by a built-in archival structure following renowned international standards (11), a user support system (10), document indexing (9), free deposit (8), research community networking features (7), data visualisation features (5), web services for audio annotation, customizable accessibility levels (4 each), web services for audio restoration (2), and archive versioning (1). Eighteen respondents opted for receiving updates on *Archivio ViVo* and subscribing to its mailing list (item 73). The last two points were about the prototype idea of OAr Journal. Item 74 asked participants if they were interested in an Italian journal on oral archives, receiving positive responses (17 “yes” vs. 5 “no”). Item 75 followed up and asked if a system equating archival efforts to acknowledged research outputs would incentivise respondents to take care of their oral documents. Similarly to the previous point, sixteen of them answered affirmatively.

3.10. Personal information regarding all the subjects involved

The last section of the census gathered information regarding all the parties involved with the archive: the respondent, the owner, the creator, whoever made the recordings, the person in charge of their conservation and dealing with potential legal issues. Moreover, it also pinpointed the actual location of the archives and clarified their level and modality of access. In the report presented below, we will gloss over the items which served the sole purpose of providing us with contact information.

Mean age of the respondents was 56, with the youngest participant being 26, and the oldest 85 years old (item 77); eleven were male, nine female, and two did not specify their sex (item 78). One respondent was a Ph.D. student, another an ex-Master of Arts student; one a post-doc researcher (*assegnista*), two fixed-term researchers, two associate professors, two full professors, one emeritus; two research enthusiasts, one independent researcher; three archive/institution/project managers; one secondary school teacher; one documentary film director and producer; three retired individuals; the role in the research community of one participant was not specified (item 79). Nine respondents stated their affiliation (item 80) to Tuscan Universities or Institutions, two were affiliated with Italian, but non-Tuscan, Universities or Institutions, and four worked abroad (the Netherlands, France, the US, Australia). Item 81 asked them about their affiliation to research associations: four of our twenty-two respondents did not reply, six did not have any research association affiliation, while the remaining twelve showcased one or more affiliations. Item 82 presented respondents with a CATA list of options and asked them to state their research field. The most commonly selected options were phonetics/phonology and dialectology (8 each), followed by oral history (6), sociolinguistics (5), anthropology (4), ethnomusicology and musicology (2 each), sociology, psychology, applied linguistics, political science, morphosyntax, neuroscience, literary studies, folklore studies, historical linguistics, film studies and history of art (1 each).

Most of the respondents were also the owners of the archive (17; item 86), their ownership mostly resulting from their being also the original creators of the archive (15 out of 17), while one form mentioned that ownership had been established through a donation, another through a transaction (item 87). Conversely, most of the non-owners specified that these rights belonged to the institution hosting the archive itself (item 88). Item 89 focused on archive creators: nineteen of our respondents considered themselves as such, while the remaining forms suggested that the creator was an institution (in two cases) or another individual researcher (item 90). A similar picture can be drawn through the answers to item 91 about who actually did the recordings. Eighteen respondents were also directly involved with the recording process; only two forms specified that this was done by other unnamed individuals (item 92).

Item 93 asked where the archives were physically located. One participant did not reply to this question. Most of them were in Tuscany (19 out of 23), especially in the *Provincia* of Florence (9), followed by Arezzo, Pistoia, Prato, Siena (2 each), Pisa, and Livorno/Massa Carrara (1 each). The other 3 archives were located in Genoa, Mechelen, and Melbourne. Item 94 then investigated whether the documents were accessible to third parties or not. Five archives (out of 23) were not accessible at all; seven were granted total accessibility, and eleven only partial accessibility. Item 95 presented these eighteen respondents with a CATA containing multiple options about the available modalities of access to the documents. The most commonly chosen modality was in-person access through prior appointment and research plan (15 out of 18), followed by online sharing upon request (3), and even free online access (2). The remaining five forms led to a similar CATA with options concerning the reasons for non-accessibility (item 96). Privacy issues was selected four times, followed by carrier fragility, lack of play-back devices, inadequate spaces for in-person appointments, and no assigned personnel (one time each); moreover, two forms contained original remarks, i.e., legal uncertainties and claims of research priority of whoever made the recordings.

Respondents were then asked about who was in charge of the preservation of the archive (item 97). Most of the forms (20 out of 23), reported that the respondents themselves took care of this aspect; the other three selected “no one”, “the owner of the archive”, and “another individual”. In this last case, it was further specified (item 98) that the archive was curated by the scientific director of a university laboratory. Lastly, the census asked who was in charge of the legal aspects of the archives (item 99). In this case, the situation was more nuanced: while 10 forms reported that the respondents themselves took care of legal issues, seven stated that no one was in charge of these, and others indicated the archive owner, other individuals, or unspecified other situations (2 each). The last item (100) contained the names of the appointed individuals.

4. DISCUSSION

In the previous section, I recounted the results of the *Grafo Reloaded* census. As was acknowledged above (§2.2), a synthesis of twenty-three archives can barely scratch the surface of the rich complexity of the oral archive situation in the region of Tuscany. For this reason, the patterns of responses to this census may result extremely local and hardly generalisable. Nonetheless, this section will showcase a selection of threads across the census with the

aim of encouraging prospective actions and verifications; at the same time, I will also keep an eye on any element which could unveil the dynamics of our success. This discussion will try to stick to the order of the census items; however, many threads feed on multiple subsections, so that the reader will find several exceptions to this principle.

Starting from the very beginning, the huge variety of analog carrier types reported in our very few forms (item 1) is both intriguing and daunting from the viewpoint of preservation. From a single cylinder¹⁰ to the vast number of audiocassettes, every carrier type has its own (more or less) specific needs (e.g., Behl 2015). The obsolescence and mutual incompatibility of play-back devices add up to what Orio et al. (2009, 202) vividly call a “Tower of Babel” concerning analog carriers, as the preservation of their reproduction is “hopeless” in the long run. For this reason, an active approach involving the digitisation of the carrier contents is strongly recommended. In our census, almost half of the described archives contain carriers with at least incipient signs of wear (item 5); however, only a single respondent stated that anything relating to analog restoration had been carried out (items 7-11) and, in two cases, carriers could not be played back due to lack of specific devices. An improvement of audio quality is seemingly not the highest priority on the part of our respondents (item 72), this perhaps having something to do with the fact that the documents in the described archives are seldom listened to (items 31-32)¹¹. Nonetheless, eleven archives underwent some digitisation process (item 18)¹²; curiously, these actions were not perceived as related to long-term preservation (item 29) and were maybe driven by an attempt to ease access to the documents. However, facilitating the access to and the distribution of oral documents is exactly what is considered the key factor of preservation in the twenty-first century (Cohen 2001, 21). This lack of awareness may have led to an underestimation of the philological importance of producing (and archiving) documentation of the digitisation process (items 21-22; see Bressan 2018). Along the same lines, digitisation was mostly performed without the intervention of third parties (item 20): indeed, the only two archives with digitisation documentation were not processed by the individuals involved with the archive. The *Vademecum* already provided Italian researchers with a list of professional digitisation centres (but see below); it is up to future actions to reach a better understanding of the reasons behind this preference for autonomous interventions.

¹⁰ Note that the presence of this single specimen of the first commercialized audio carrier in one of our forms (a Florentine archive) is slightly problematic. In fact, the researcher stated that the earliest recordings contained in the archive were made in 1973 (item 33); however, in specific contexts, cylinders were used up to the Sixties (Shambarger 1995; for an example of later applications, see e.g., Williams 2018). Nonetheless, the respondent stressed the existence of this carrier in item 3, where he provided a quantification of this type only. As cylinders were previously found in a single Tuscan archive in Siena (*Fondazione Siena Jazz*; Cappelli and Rioda 2009), this point may warrant further investigations. The same goes for the single mention to magnetic wire recordings (see Daniel, Mee, and Clark 1999 for historical remarks), which were previously found in the large archive of *Nomadelfia* (Grosseto) by Andreini and Clemente (2007), but not by Cappelli and Rioda (2009). Our testimony comes from a private archive stored in two private homes between Livorno and the Massa Carrara Province; however, the respondent talked about these carriers in item 1 only, so that this information should be properly verified.

¹¹ Since most of the archives were the object of some form of publication (item 61), one may ask if this low frequency of listening to oral documents is connected to a general avoidance of multiple publications from the same data or if the archives already used up all their publication potential at the time of the census. As the issue is quite controversial in experimental research (Kirkman and Chen 2011), asking participants whether they ever publish multiple works based on the same oral documents or not could be of interest to future surveys.

¹² Note that, in comparison, our percentage is higher than that presented in Cappelli and Rioda (2009), in which 85% of the institutions did not digitise their materials.

Coming to the raw quantity of carriers (item 2), the numbers of the *Grafo Reloaded* census (approximately 4,417) are understandably lower than those of Andreini and Clemente (2007) (115,072) and Cappelli and Rioda (2009) (74,616), but are higher than those of the original *Grafo* project (1,971), despite our fewer respondents. The original *Grafo* project conceptualised the size of an archive (as it appears through access portals) as the number of contained documents (i.e., interviews, narratives, etc.) rather than the number of analog carriers. Significantly enough, the carrier numbers commented here can be found in the presentation¹³ precluding to Calamai et al. (2018), but not in the actual paper itself, where the above-described view is generally endorsed. Elsewhere (e.g., Biliotti and Calamai 2019), the number of documents (2,208) is reported in project overviews. Extensive data on the carrier/document ratio could be of great interest not only to achieve an overview of the recording practices of researchers of different times and dealing with different types of equipment, but also to have a more precise estimate of the interpretative effort needed to process an oral archive in compliance with this conceptualisation. As a census cannot realistically ask respondents themselves to process huge quantities of materials in order to fill in a form, this will hopefully be obtainable, for example, after the release and subsequent public use of the *Archivio Vi.Vo.* platform, which has been specifically designed in line with these principles. Returning to the differences in carrier quantities, it should be recalled that the *Grafo* census was expedient to the digitisation and archival processing of the materials, so that their numbers reflect the accomplishments of a complex pipeline over years of project work. Be that as it may, our census managed to highlight medium-sized institutional archives (such as those at the CGIL) which were not covered by the previous similar regional endeavours. All in all, it might be suggested that the existing Tuscan censuses are not working as an update chain but are slowly and additionally enlarging our pool of known archives in a horizontal way. If that is the case, the idea of a participative permanent platform suitable for the autonomous reporting of oral archives may indeed be seen as an efficient solution for saving time and resources (Michetti 2022). Nonetheless, diachronic concerns still remain, as we may only provocatively assume that the private archives, whose existence has been reported only once and are not the object of official declarations of importance, have not vanished into thin air.

The section of the census on digital documents was probably the most perplexing. As was previously mentioned (§2.2), the *Grafo Reloaded* census put great emphasis on the discovery of digital-native archives. This expectation was bound to various general considerations. Technological advancements make audio recording an increasingly easy, accessible, and ubiquitous activity (for oral history, see e.g. Sielaff 2024; similar statements can be found outside the disciplines dealing with the human voice – Pavan et al. 2022). As a consequence, recordings become an integral part of everyday life even outside of specific research activities, and it comes as no surprise that scholars are more and more interested in personal digital archiving (Allegrezza 2023): indeed, even our census managed to describe an archival effort driven by an attempt to make family memories (item 40). Moreover, there is also a generational element at play here, in that young researchers starting to make their own recordings are explicitly taught to opt out the use of analog devices (e.g., De Decker and Nycz 2018, 124), so that future oral archives will presumably be, for the most part, digital only. Overall, all previous

¹³ <https://www.clarin.eu/sites/default/files/4-oral-session-2-2-calamai.pdf>, accessed July 1, 2024.

censuses were secondarily concerned with digital files, so that digital-native archives seemed to us a relatively unexplored field of inquiry. The census results did not reflect this contemporary trend, as digital-only archives were still outnumbered by mixed and analog-only ones. Of course, this information may depend on idiosyncratic aspects of the small respondent pool; moreover, since most of these archives were indeed described by young researchers, a census more specifically targeting this demographic subset should be able to provide a better view on digital archiving practices involving oral materials. Just to make sure that terminological misunderstandings are not a factor, future endeavours should probably explicitly state from the beginning that descriptions of digital recordings are allowed. Since popular stereotypes about archives focus on oldness and physical materials and are at odds with the digital age (see Patterson 2016 for a literature recollection and recent data), potential respondents with digital archives may have erroneously thought that the census was not for them.

Most of the archives contain uncompressed .wav digital recordings (item 14). General consensus considers this format as the standard for audio preservation (e.g., Prentice 2016, 134); moreover, uncompressed formats are needed by researchers dealing with acoustic analysis (De Decker and Nycz 2018, 124-5). This may be a sign of a good level of technical awareness, which is indeed reflected in the positive responses to item 63 on previous technical knowledge. On a side note, compared to the previous section, the open text items on digital files requiring specific quantifications and checks (15, 17) received relatively less accurate responses. This has probably little to do with the nature of the tasks themselves, since they can be mostly accomplished through automatic searches. Conversely, it may indicate that survey fatigue was already in play after the first section: in support of this view, a later similar item on other kinds of documents (47) was left almost deserted. Indeed, research has shown that open ended questions are particularly susceptible to time-driven fatigue, so that earlier items of this kind are expected to be more complete. Future censuses which are primarily concerned with precise quantification should try to move related items to the top of their forms (Jeong et al. 2023).

Only six archives were (at least partially) uploaded online, and none of them in a research repository (items 23-24), despite the fact that many of the respondents declared interest in doing so (item 71). While item 72 lists many technical features suitable for incentivising similar behaviours, we may only guess at why this has not already been done. Accessibility does not seem to be the core issue here, as the majority of respondents stated that their materials could be at least partially consulted following specific procedures (items 94-95)¹⁴. On the contrary, as was already noted through an analysis of personal communications (Piccardi and Calamai Forthcoming; §2.2), legal insecurity evidently crept into the census. Indeed, this was at least partially due to objective reasons: only eight forms state that consent forms signed by recorded individuals are available (item 58) and, more generally, only one researcher had any legal knowledge prior to making the recordings (item

¹⁴ This information clashes with the widespread inaccessibility lamented in Andreini and Clemente (2007) and, more recently, in Galatà and Calamai (2019). Conversely, the institutional archives described in Cappelli and Rioda (2009) are for the most part accessible. Our results may either signal a positive historical trend, or just be biased by the type of form: researchers who were willing to fill in an extensive census like ours were probably also already open to let others know about their materials and use them. The two forms declaring open internet accessibility referred to uploads on YouTube and a dedicated website.

67). Moreover, items about legal issues saw noticeable numbers of “Not Available” (NA) responses, implying uncertainties or unwillingness to talk about core issues (such as copyright, item 54, which received a whopping 15 total number of NA responses). In item 99, seven respondents stressed that no one was responsible for the legal aspects of the archive. Of course, in this respect, researchers can count on resources for individual study – the *Vademecum* has an extensive section on legal issues and several papers describe problems and solutions in the processing of specific oral archives (e.g., Calamai et al. 2018). However, hasty analogical reasoning can be risky when it comes to legal issues (see Sunstein 2021, 20-3 for discussion), and researchers may feel the need for more opportunities for in-depth learning. As in the case of the only respondent claiming prior legal knowledge, dedicated workshops might do the trick. More ambitiously speaking, given the importance of university courses in acquiring other kinds of knowledge (item 64), these learning opportunities might be structurally linked to degree courses in the humanities.

The same goes for archival issues. In section 3, participants were asked about aspects of archival description (items 27-28). While most of the archives underwent some kind of description, it should be noted that participants were not particularly willing to share their methods; in general, none of the archives, whether private or institutional, were described using any archival standard model. Standards are primarily used to achieve description consistency (e.g., Millar 2017, 103), which leads to ease of information exchange¹⁵. As we mentioned above, in the digital age, this is synonymous with longevity (Lorist and Van der Meer 2001). The census results revealed that researchers involved with recordings are usually not trained in archival issues (item 65). According to Allegrezza (2019), the abundance of personal data digitally produced by everyone fosters the popular need for basic archival literacy. In Italy, archival education is undergoing major changes (see the contribution by Manzi and Stamuli in this issue). If nothing else, these results suggest that specific actions should be planned to give researchers in other disciplines the opportunity to acquire appropriate specific knowledge concerning the systematisation and consequent preservation of their materials.

Section four of the census highlighted a huge variety of responses in such a small number of forms. From the Fifties to the present day, from the big cities to the hamlets of every province, reflecting disparate research methods and rationales, and concerning a very diverse pool of languages, spoken text type and themes, oral archives in Tuscany are definitely confirmed as being a treasure trove of information (Calamai and Biliotti 2017b). In our census, the province of Florence was the most frequently mentioned as a setting of sound recordings (item 34). This patterns well with the fact that the Florentine area is home to the highest number of oral archives (item 93), as has been known since Andreini and Clemente (2007). Oral history and phonetics/phonology were highly represented disciplines, both at the archive (item 35) and the respondent (item 82) level. These disciplines were probably over-represented because of our closer social proximity to the respective research communities (Piccardi and Calamai Forthcoming; see also the Editorial); possibly connected to this issue, on the other hand, is the fact that archives of psychological interest are completely absent,

¹⁵ In the form, the *General International Standard Archival Description* (ISAD G) was mentioned as an example. It can be argued that the CHAT model followed by one respondent (§3.5) contributes to this aspect at some level; in this regard, see Calamai and Frontini (2018).

and the only researcher mentioning this area of research was in actual fact a linguist. Psychologists have a long history of using audio(video) recordings to objectively capture the contents of therapy sessions (see the literature review in Briggie et al. 2016), requiring a particularly careful approach from the ethical and legal side (Funkenstein, Kessler, and Schen 2014). Moreover, concerns about the hosting and sharing of audio recorded for experimental studies are featured in their current research discussions (Gilmore, Kennedy, and Adolph 2018). It is unclear whether these documents are considered by the respective research communities as fitting into the definition of “oral archive” promoted through our census. Future inquiries specifically targeting the national psychology community could be a valuable addition to the research line on oral archives in Italy.

Other issues related to the two last sections of the census (9 and 10) have already been commented in Piccardi and Calamai (Forthcoming). Here, suffice it to say that, despite the overall positive responses to the declaration of interest items, the *Vademecum* proved to be unknown to most of our respondents, as only four had heard of it prior to the census (item 69). This information outlines a still unsatisfactory level of diffusion of the tool conceived for the lowering of knowledge barriers in the curation of oral archives in Italy (§2.2). Thus, a periodic monitoring of its use in different research communities is strongly advisable: this can be done, for example, through sectorial literature reviews and individual reporting of actions centred around the volume, consequently allowing for more tailored future dissemination strategies. In the meantime, to end this chapter on a hopeful note, we were delighted to see that the concept of OAr Journal and its intents caught the interest of many of our respondents (items 74-75).

5. CONCLUSIONS. ON NEGATIVITY’S ULTIMATE END

This paper dealt with a story of unsuccess which holds a foundational value for OAr Journal. The meagre public response described in these pages led us to see the results of the *Grafo Reloaded* project less as a census and more a lesson. After examining the outcomes of a report on oral archives in Ireland, Ferriter (1998, 95) stated that “it is difficult to know where to go from here”. In this regard, the data examined in this paper provides at least some directions. We know that oral archives in Tuscany represent a source of irreplaceable information for many disciplines dealing with human voice. In order to fully unlock their potential, we need to raise awareness on what preservation actually means in the twenty-first century, and on the philological importance of many technological processes which are often deemed unrelated to it. We need to soothe the rampaging risk aversion of researchers dealing with oral data, who are increasingly surrounded by legal regulations, ethical checks, and project requirements without having access to adequate learning opportunities on the subject. We need to begin to acknowledge that archival dams are plausible containments of data floods, and that archival science can no longer be considered a self-contained universe. We need to reach out to other research communities and hear their perspectives, while devising dissemination strategies of existing resources and new incentives for the curation of oral archives.

Circling back to the beginning of this ride, together with Casellato and Lampe (2023), it may be argued that all this cannot be achieved without a radical destigmatisation of the negative aspects of science, whether they be missteps, subpar results, overly ambitious goals,

fears or misconceptions. In this regard, OAr Journal makes a first step, and not just for the sake of disciplinary advancements. Endorsing good practices, fairness, and several virtuous checklists cannot by any means become a new facet of the harmful illusion of superhumanity in contemporary academia (Horton 2020, 3). On the contrary, these risk becoming a sterile façade without a firm promotion of the inherent tentativeness of the scientific process leading to them and their application. Most importantly, if failure is bound to happen because of a physical law (Firestein 2016), being open to it means defending the human nature of the research process; it means, above all, defending honesty in science (Mehta 2019).

ACKNOWLEDGMENTS

This work has been possible thanks to the financial support of the *Gra.fo Reloaded* project, as described above in §2.1. A census is, by definition, something that requires the cooperation of a lot of people. Here I would like to thank Maria Francesca Stamuli and Giovanni Michetti for their invaluable suggestions during the development of the form, all the kind participants in the pre-test phase for their attentive examination of the census prototype, and the dedicated personnel of our partners for their help in the dissemination of the census. Of course, my deepest gratitude goes to all the respondents to the actual census. I hope that this work was enough to give credit to the time and effort you dedicated to this project. Other people and institutions contacted me during this year and a half, but then opted out of the census: our conversations were of great help to the conceptualization of this journal, so that I would like to express my gratitude to you as well. Thanks also to Alberto Zhang, and I am sorry that I did not find a solution to make the most out of the materials he sent me. Lastly, I deeply appreciate the kind participation of Temistocle Franceschi and Petra Vangelisti in our pilot interviews. Leaving aside the census endeavour, I am especially grateful to the Data Protection Officer of the University of Siena, Chiara Silvia Armida Angiolini, who guided us through all the legal conundrums of *Gra.fo Reloaded* with a friendly eye full of sincere interest. As was made clear in the main text of this article, Silvia Calamai supervised the whole *Gra.fo Reloaded* project, all the aspects of this work being constantly discussed with her.

REFERENCES

- Allchin, Douglas. 2012. "Teaching the Nature of Science through Scientific Errors". *Science Education*, 96(5): 904-26. <https://doi.org/10.1002/sce.21019>.
- Allegrezza, Stefano. 2019. "The Importance of Personal and Family Digital Archives for a New Perception of the Archival Discipline". *Atlanti*, 29(2): 58-76.
- Allegrezza, Stefano, ed. 2023. *Personal Digital Memories: i fondi di persona dall'analogico al digitale*. Torre del Lago Puccini: Civita Editoriale.
- Andreini, Alessandro, and Pietro Clemente, eds. 2007. *I custodi delle voci. Archivi orali in Toscana: primo censimento*. Florence: Regione Toscana.
- Atkeson, Lonna Rae, and Alex N. Adams. 2018. "Mixing Survey Modes and Its Implications". In *The Oxford Handbook of Polling and Survey Methods*, edited by Lonna Rae Atkeson and R. Michael Alvarez, 53-75. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190213299.013.35>.

- Barrera, Giulia, Alfredo Martini, and Antonella Mulè, eds. 1993. *Fonti orali. Censimento degli istituti di conservazione*. Rome: Ministero per i Beni Culturali e Ambientali, Ufficio Centrale per i Beni Archivistici.
- Barwich, Ann-Sophie. 2019. "The Value of Failure in Science: The Story of Grandmother Cells in Neuroscience". *Frontiers in Neuroscience*, 13. <https://doi.org/10.3389/fnins.2019.01121>.
- Behl, Harrison. 2015. "Audio Formats: Characteristics and Deterioration". In *ARSC Guide to Audio Preservation*, edited by Sam Brylawski, Maya Lerman, Robin Pike, and Kathlin Smith, 14-36. Eugene / Washington: Association for Recorded Sound Collections / Council on Library and Information Resources / The Library of Congress.
- Benedetti, Amedeo. 2002. *Gli archivi sonori: fonoteche, nastroteche e biblioteche musicali in Italia*. Genoa: Erga.
- Biliotti, Francesca, and Silvia Calamai. 2019. "Web Dissemination of Tuscan Oral Archives". *Bulletin de l'AFAS. Sonorités*, 45. <https://doi.org/10.4000/afas.3439>.
- Bradley, Simon. 2016. "Archaeology of the Voice: Explaining Oral History, Locative Media, Audio Walks, and Sound Art as Site-Specific Displacement Activities". PhD diss. University of Huddersfield.
- Bressan, Federica. 2018. "A Philological Approach to Sound Preservation". In *Research Methods for the Digital Humanities*, edited by Lewis Levenberg, Tai Neilson, and David Rheams, 243-61. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-96713-4_14.
- Briggie, Alexis M., Mark J. Hilsenroth, Francine Conway, J. Christopher Muran, and Jonathan M. Jackson. 2016. "Patient Comfort with Audio or Video Recording of Their Psychotherapy Sessions: Relation to Symptomatology, Treatment Refusal, Duration, and Outcome". *Professional Psychology: Research and Practice*, 47(1), 66-76. <https://doi.org/10.1037/a0040063>.
- Buonomini, Anna. 1987. "Sul Cappel d'Orlando e su quanto si ricorda dei Reali di Francia in alcune località dell'Appennino Tosco-Emiliano. Repertori di tradizione orale". In *Sulle orme di Orlando: leggende e luoghi carolingi in Italia. I paladini di Francia nelle tradizioni italiane: una proposta storico antropologica*, edited by Anna I. Galletti and Roberto Roda, 159-65. Padua: Interbooks.
- Calamai, Silvia, and Francesca Biliotti. 2017a. "The *Gra.fo* Project: From Collection to Dissemination". *Umanistica Digitale*, 1: 85-103. <https://doi.org/10.6092/issn.2532-8816/7230>.
- Calamai, Silvia, and Francesca Biliotti. 2017b. "Archivi orali e migrazione. La costruzione del racconto e il repertorio verbale toscano". *Mnemosyne o la costruzione del senso*, 10: 99-116.
- Calamai, Silvia, and Francesca Frontini. 2016. "Not Quite your Usual Kind of Resource. *Gra.fo* and the Documentation of Oral Archives in CLARIN". In *CLARIN Annual Conference Proceedings 2016*.
- Calamai, Silvia, and Francesca Frontini. 2018. "FAIR Data Principles and Their Application to Speech and Oral Archives". *Journal of New Music Research*, 47(4), 339-54. <https://doi.org/10.1080/09298215.2018.1473449>.
- Calamai, Silvia, Chiara Kolletzek, Aleksei Kelli, and Francesca Biliotti. 2018. "Authorship and Copyright Ownership in the Digital Oral Archives Domain: The *Gra.fo* Digital Archive in the CLARIN-IT Repository". In *Selected papers from the CLARIN Annual Conference 2017*, edited by Maciej Piasecki, 112-27. Linköping: Linköping University Electronic Press.
- 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/Boston: De Gruyter. <https://doi.org/10.1515/9783110767377-025>.
- Calamai, Silvia. 2011. "Ordinare archivi sonori: il progetto *Gra.fo*". *Rivista italiana di dialettologia*, 35: 135-64.
- Cappelli, Fabrizio, and Antonella Rioda. 2009. "Archivi sonori in Toscana: un'indagine". *Musica/Tecnologia*, 3: 9-69.

- Casellato, Alessandro, and Urška Lampe. 2023. "Learning from Mistakes, Crossing Boundaries. Oral History between Italy and Slovenia". *Acta Histriae*, 31(3): 372-76.
- Cohen, Elizabeth. 2001. "Preservation of Audio". In *Folk Heritage Collections in Crisis*, 20-27. Washington: Council on Library and Information Resources.
- Daniel, Eric D., C. Denis Mee, and Mark H. Clark, eds. 1999. *Magnetic Recording: The First 100 Years*. Piscataway: IEEE Press.
- De Decker, Paul, and Jennifer Nycz. 2018. "The Technology of Conducting Sociolinguistic Interviews". In *Data Collection in Sociolinguistics. Methods and Applications. Second Edition*, edited by Christine Mallinson, Becky Childs, and Gerard Van Herk, 123-30. New York/London: Routledge.
- Dickersin, Kay. 2005. "Publication Bias: Recognizing the Problem, Understanding Its Origins and Scope, and Preventing Harm". In *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, edited by Hannah R. Rothstein, Alexander J. Sutton, and Michael Borenstein, 11-33. Chichester: John Wiley & Sons. <https://doi.org/10.1002/0470870168.ch2>.
- Eddington, David. 2008. "Linguistics and the Scientific Method". *Southwest Journal of Linguistics*, 27(2): 1-16.
- Fanelli, Daniele. 2012. "Negative Results are Disappearing from Most Disciplines and Countries". *Scientometrics*, 90(3): 891-904. <https://doi.org/10.1007/s11192-011-0494-7>.
- Ferriter, Diarmaid. 1998. "Oral Archives in Ireland: A Preliminary Report". *Irish Economic and Social History*, 25(1): 91-95. <https://doi.org/10.1177/033248939802500108>.
- Firestein, Stuart. 2016. *Failure. Why Science is So Successful*. Oxford: Oxford University Press.
- Fitzgerald, Colleen M. 2021. "A Framework for Language Revitalization and Documentation". *Language*, 97(1): 1-11. <https://doi.org/10.1353/lan.2021.0006>.
- Franco, Annie, Neil Malhotra, and Gabor Simonovits. 2014. "Publication Bias in the Social Sciences: Unlocking the File Drawer". *Science*, 345(6203): 1502-5. <https://doi.org/10.1126/science.1255484>.
- Funkenstein, Amy B., Katherine A. Kessler, and Cathy R. Shen. 2014. "Learning Through the Lens: Ethical Considerations in Videotaping Psychotherapy". *Harvard Review of Psychiatry*, 22(5): 316-22. <https://doi.org/10.1097/HRP.0000000000000024>.
- Galatà, Vincenzo, and Silvia Calamai. 2019. "Looking for Hidden Speech Archives in Italian Institutions". In *Selected Papers from the CLARIN Annual Conference 2018*, edited by Inguna Skadiņa and Maria Eskevich, 46-55. Linköping: Linköping University Electronic Press.
- Gilmore, Rick O., Joy Lorenzo Kennedy, and Karen E. Adolph. 2018. "Practical Solutions for Sharing Data and Materials from Psychological Research". *Advances in Methods and Practices in Psychological Science*, 1(1), 121-30. <https://doi.org/10.1177/2515245917746500>.
- Gomes, Dylan G. E., Patrice Pottier, Robert Crystal-Ornelas, Emma J. Hudgins, Vivienne Foroughi-rad, Luna L. Sánchez-Reyes, Rachel Turba, Paula Andrea Martinez, David Moreau, Michael G. Bertram, Cooper A. Smout and Kaitlyn M. Gaynor. 2022. "Why Don't We Share Data and Code? Perceived Barriers and Benefits to Public Archiving Practices". *Proceedings of the Royal Society B*, 289(1987). <https://doi.org/10.1098/rspb.2022.1113>.
- Grimaccia, Elena, Alessia Naccarato, Gerardo Gallo, Novella Cecconi, and Alessandro Fratoni. 2023. "Characteristics of Respondents to Web-Based or Traditional Interviews in Mixed-Mode Surveys. Evidence from the Italian Permanent Population Census". *Journal of Official Statistics*, 39(1): 1-26. <https://dx.doi.org/10.2478/JOS-2023-0001>.
- Hardin, Russell. 2006. *Trust*. Cambridge/Malden: Polity Press.
- Horton, John. 2020. "Failure Failure Failure Failure Failure Failure: Six Types of Failure Within the Neoliberal Academy". *Emotion, Space and Society*, 35(Special Issue on Reclaiming Failure in Geography: Academic Honesty in a Neoliberal World). <https://doi.org/10.1016/j.emospa.2020.100672>.
- Isbell, Daniel R., Dan Brown, Meishan Chen, Deirdre J. Derrick, Romy Ghanem, María Nelly Gutiérrez Arvizu, Erin Schnur, Meixiu Zhang, and Luke Plonsky. 2022. "Misconduct and Questionable

- Research Practices: The Ethics of Quantitative Data Handling and Reporting in Applied Linguistics". *The Modern Language Journal*, 106(1), 172-95. <https://doi.org/10.1111/modl.12760>.
- Jeong, Dahyeon, Shilpa Aggarwal, Jonathan Robinson, Naresh Kumar, Alan Spearot, and David Sung-ho Park. 2023. "Exhaustive or Exhausting? Evidence on Respondent Fatigue in Long Surveys". *Journal of Development Economics*, 161. <https://doi.org/10.1016/j.jdeveco.2022.102992>.
- Kirkman, Bradley L., and Gilad Chen. 2011. "Maximizing Your Data or Data Slicing? Recommendations for Managing Multiple Submissions from the Same Dataset". *Management and Organization Review*, 7(3): 433-46. <https://doi.org/10.1111/j.1740-8784.2011.00228.x>.
- Kortmann, Bernd. 2021. "Reflecting on the Quantitative Turn in Linguistics". *Linguistics*, 59(5): 1207-26. <https://doi.org/10.1515/ling-2019-0046>.
- Lipparini, Giuseppe. 1930. *I racconti di Cutigliano*. Milan: Mondadori.
- Liu, Mingnan, and Laura Wronski. 2018. "Examining Completion Rates in Web Surveys via Over 25,000 Real-World Surveys". *Social Science Computer Review*, 36(1): 116-24. <https://doi.org/10.1177/0894439317695581>.
- Lorist, Jeroen, and Kees van der Meer. 2001. "Standards for Digital Libraries and Archives: Digital Longevity". In *New Developments in Digital Libraries, Proceedings of the 1st International Workshop on New Developments in Digital Libraries*, NDDL 2001, In conjunction with ICEIS 2001, Setúbal, Portugal, July 6-7, 2001.
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk. Third Edition*. Mahwah: Lawrence Erlbaum Associates.
- Mehta, Devang. 2019. "Highlight Negative Results to Improve Science". *Nature Career Column*. <https://doi.org/10.1038/d41586-019-02960-3>.
- Michetti, Giovanni. 2022. "Il censimento degli archivi sonori in Italia". Paper presented at Un anno di 'Vademecum per il trattamento delle fonti orali', Rome, October 27, 2022.
- Millar, Laura A. 2017. *Archives. Principles and Practices. Second Edition*. London: Facet Publishing.
- Nicholls, Helen, Matthew Nicholls, Sahra Tekin, Danielle Lamb, and Jo Billings. 2022. "The Impact of Working in Academia on Researchers' Mental Health and Wellbeing: A Systematic Review and Qualitative Metasynthesis". *Plos One*, 17(5). <https://doi.org/10.1371/journal.pone.0268890>.
- Nichols, Elizabeth, Erica Olmsted-Hawala, Temika Holland, and Amy Anderson Riemer. 2020. "Usability Testing Online Questionnaires: Experiences at the U.S. Census Bureau". In *Advances in Questionnaire Design, Development, Evaluation and Testing*, edited by Paul Beatty, Debbie Collins, Lyn Kaye, Jose Luis Padilla, Gordon Willis, and Amanda Wilmot, 315-48. Hoboken: John Wiley. <https://doi.org/10.1002/9781119263685.ch13>.
- Orio, Nicola, Lauro Snidaro, Sergio Canazza, and Gian Luca Foresti. 2009. "Methodologies and Tools for Audio Digital Archives". *International Journal on Digital Libraries*, 10: 201--20. <https://doi.org/10.1007/s00799-010-0060-6>.
- Parkes, Eileen. 2019. "Scientific Progress is Built on Failure". *Nature Career Column*. <https://doi.org/10.1038/d41586-019-00107-y>.
- Patterson, Caitlin. 2016. "Perceptions and Understandings of Archives in the Digital Age". *The American Archivist*, 79(2): 339-70. <https://doi.org/10.17723/0360-9081-79.2.339>.
- Pavan, Gianni, Gregory Budney, Holger Klinck, Hervé Glotin, Dena J. Clink, and Jeanette A. Thomas. 2022. "History of Sound Recording and Analysis Equipment". In *Exploring Animal Behavior Through Sound: Volume 1. Methods*, edited by Christine Erbe and Jeanette A. Thomas, 1-36. Cham: Springer. https://doi.org/10.1007/978-3-030-97540-1_1.
- Pfeffer, Christian, and Bjorn R. Olsen. 2002. "Editorial: Journal of Negative Results in Biomedicine". *Journal of Negative Results in Biomedicine*, 1. <https://doi.org/10.1186/1477-5751-1-2>.
- Piccardi, Duccio, and Silvia Calamai. 2023. "How Many Oral Archives are in your Home? Piloting a New Tuscan Census in the *Grafo Reloaded* Project". In *La memoria digitale. Forme del testo e*

- organizzazione della conoscenza. Atti del XII convegno annuale AIUCD*, edited by Emmanuela Carbé, Gabriele Lo Piccolo, Alessia Valenti, and Francesco Stella, 389-94. University of Siena: AIUCD.
- Piccardi, Duccio, and Silvia Calamai. Forthcoming. "Fear of FAIR? Towards a new Italian incentive to oral data curation".
- Popoff, Martin. 2009. *Goldmine Record Album Price Guide. 6th Edition*. Iola: Krause.
- Pozzebon, Alessandro, Francesca Biliotti, and Silvia Calamai. 2016. "Places Speaking with Their Own Voices. A Case Study from the *Grafo* Archives". In *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection. EuroMed 2016*, edited by Marinos Ioannides, Eleanor Fink, Antonia Moropoulou, Monika Hagedorn-Saupe, Antonella Fresa, Gunnar Liestøl, Vlatka Rajcic, and Pierre Grussenmeyer, 232-9. Cham: Springer. https://doi.org/10.1007/978-3-319-48974-2_26.
- Prentice, Will. 2016. "Digital Preservation of Audio Content". In *Managing Digital Cultural Objects. Analysis, Discovery and Retrieval*, edited by Allen Foster and Pauline Rafferty, 129-39. London: Facet Publishing. <https://doi.org/10.29085/9781783301539.007>.
- Rao, Gautam, and David Mingay. 2001. "Report on Usability Testing of Census Bureau's Dynamaps CD-ROM Product". <https://web.archive.org/web/20230605040654/https://infovis.cs.vt.edu/oldsite/cs5764/papers/dynamapsUsability.pdf>.
- Revilla, Melanie, and Carlos Ochoa. 2017. "Ideal and Maximum Length for a Web Survey". In *International Journal of Market Research*, 59(5): 557-65. <https://doi.org/10.2501/IJMR-2017-039>.
- Roettger, Timo. B. 2021. "Preregistration in Experimental Linguistics: Applications, Challenges, and Limitations". *Linguistics*, 59(5): 1227-49. <https://doi.org/10.1515/ling-2019-0048>.
- Rogers, Anna, João Sedoc, and Anna Rumshisky. 2020. *Proceedings of the First Workshop on Insights from Negative Results in NLP*. Association for Computational Linguistics.
- Sammut, Roberta, Odette Griscti, and Ian J. Norman. 2021. "Strategies to Improve Response Rates to Web Surveys: A Literature Review". *International Journal of Nursing Studies*, 123. <https://doi.org/10.1016/j.ijnurstu.2021.104058>.
- Schickore, Jutta. 2005. "'Through Thousands of Errors We Reach the Truth'—but How? On the Epistemic Roles of Error in Scientific Practice". *Studies in History and Philosophy of Science*, 36(3): 539-56. <https://doi.org/10.1016/j.shpsa.2005.06.011>.
- Schickore, Jutta. 2021. "Is 'Failing Well' a Sign of Scientific Virtue?". In *Science, Technology, and Virtues. Contemporary Perspectives*, edited by Emanuele Ratti and Thomas A. Stapleford, 178-99. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190081713.003.0011>.
- Shambarger, Peter. 1995. "Cylinder Records: An Overview". *ARSC Journal*, 26(2): 133-61.
- Shur-Ofry, Michal. 2016. "Access-to-Error". *Cardozo Arts & Entertainment Law Journal*, 34(2): 357-400.
- Sielaff, Steven. 2024. "Remote Recording Technology". In *Oral History at a Distance*, edited by Steven Sielaff, Stephen M. Sloan, Adrienne A. Cain Darough, and Michelle Holland, 106-29. New York: Routledge. <https://doi.org/10.4324/9781003206606-7>.
- Simonetti, Valentina. 2007. "Il Censimento nell'analisi archivistica e alcune considerazioni sulle fonti orali". In *I custodi delle voci. Archivi orali in Toscana: primo censimento*, edited by Alessandro Andreini and Pietro Clemente, 271-84. Florence: Regione Toscana.
- Statistics New Zealand. 2018. *2018 Census. Design of Forms*. Wellington: Stats NZ Tauranga Aotearoa. <https://www.stats.govt.nz/assets/Reports/2018-census-design-of-forms/2018-Census-Design-of-forms.pdf>.
- Stracke, Christian M. 2020. "Open Science and Radical Solutions for Diversity, Equity and Quality in Research: A Literature Review of Different Research Schools, Philosophies and Frameworks and Their Potential Impact on Science and Education". In *Radical Solutions and Open Science*, edited by Daniel Burgos, 17-37. Singapore: Springer. https://doi.org/10.1007/978-981-15-4276-3_2.

- Sunstein, Cass R. 2021. "Analogical Reasoning". *Harvard Public Law Working Paper*, 21-39. <https://doi.org/10.2139/ssrn.3938546>.
- Tavolo permanente per le fonti orali. 2023. *Vademecum per il trattamento delle fonti orali*. Rome: Ministero della Cultura, Direzione generale Archivi.
- Thönes, Sven, and Kurt Stocker. 2019. "A Standard Conceptual Framework for the Study of Subjective Time". *Consciousness and Cognition*, 71: 114-22. <https://doi.org/10.1016/j.concog.2019.04.004>.
- 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.
- Van Lent, Marlies, John Overbeke, and Henk Jan Out. 2014. "Role of Editorial and Peer Review Processes in Publication Bias: Analysis of Drug Trials Submitted to Eight Medical Journals". *Plos One*, 9(8), <https://doi.org/10.1371/journal.pone.0104846>.
- Vannette, David L., and Jon A. Krosnick. 2018. "Introduction". In *The Palgrave Handbook of Survey Research*, edited by David. L. Vannette and Jon A. Krosnick, XXI-XXIV. Cham: Palgrave Macmillan.
- Verbuyst, Rafael, and Anna Milena Galazka. 2023. "Introducing 'Navigating Failure in Ethnography': A Forum About Failure in Ethnographic Research". *Journal of Organizational Ethnography*, 12(1): 61-75. <https://doi.org/10.1108/JOE-10-2022-0027>.
- Villar, Ana, Mario Callegaro, and Yongwei Yang. 2013. "Where Am I? A Meta-Analysis of Experiments on the Effects of Progress Indicators for Web Surveys". *Social Science Computer Review*, 31(6): 744-62. <https://doi.org/10.1177/0894439313497468>.
- Visconti, Arianna. 2019. "The Reform of Italian Law on Cultural Property Export and its Implications for the "Definitional Debate": Closing the Gap with the European Union Approach or Cosmetics? Some Systemic Considerations from Criminal Law Perspective". *Santander Art and Culture Law Review*, 5(2): 159-86. <https://doi.org/10.4467/2450050XSNR.19.018.11566>.
- Wen, Ju, and Lei Lei. 2022. "Linguistic Positivity Bias in Academic Writing: A Large-Scale Diachronic Study in Life Sciences Across 50 Years". *Applied Linguistics*, 43(2): 340-64. <https://doi.org/10.1093/applin/amab037>.
- Williams, Alan. 2018. "Weapons of Mass Deception: The Invention and Reinvention of Recording Studio Mythology". In *Critical Approaches to the Production of Music and Sound*, edited by Samantha Bennett and Eliot Bates, 157-74. London/New York: Bloomsbury Academic.

APPENDIX

In this appendix, the reader will find the *Grafo Reloaded* census form in Italian, as it was originally distributed. See §3 for English comments on each item. Different options are reported in single lines and divided by slashes. The “Altro...” label implies that the respondents were able to specify other options. ND is the Italian equivalent of NA. The Google Forms skip logic paths will also be reproduced here inside angle brackets. Recall that CATA stands for “Check All That Applies”; FC means “Forced Choice”; OTQ is “Open Text Question”. [HT] signals the presence of a link to an explicative webpage. The first page contained the informed consent form.

[P2] Consistenza dell’archivio fisico¹⁶

1 [CATA] – Quali supporti fisici sono contenuti nell’archivio?

Cilindri / Dischi a 78 giri (prima del ’25) / Dischi a 78 giri (dopo il ’25) / Dischi a 45 giri / Dischi a 33 giri / Altri dischi / Filo magnetico / Compact cassette - audiocassette / Bobine a un quarto di pollice / Bobine a mezzo pollice / Bobine a un pollice / Bobine a due pollici / Supporti magnetici in altri formati / CD AUDIO / CD ROM / DAT / DVD / Hard disk interno / Hard disk esterno / SSD interno / SSD esterno / Scheda SD / Flash Drive USB / Altro...

2 [OTQ] – Quanti supporti fisici sono contenuti nell’archivio in totale? (NB: non singoli documenti)

3 [OTQ] – Sai quantificare il numero di supporti per tipologia? Compila il riquadro qui sotto nel formato [Tipologia: numero supporti, ad es. Cilindri: 3]

4 [FC] – Sono previsti incrementi del numero dei supporti fisici?

Sì / No / ND

5 [FC] – Qual è lo stato di conservazione dei supporti fisici in riferimento alla spiegazione riportata tra parentesi quadre?

Scarso [supporti fisici danneggiati che impediscono la riproduzione audio o la rendono inintelligibile a causa, ad es., di distorsioni e continui salti] / Buono [evidenti imperfezioni del supporto fisico che hanno un impatto notevole sulla riproduzione audio, ad es. rumore evidente, sibilo, crepitio] / Molto buono [evidenti imperfezioni del supporto fisico che hanno un impatto di piccola entità sulla riproduzione audio, ad es. rumore che non sovrasta il livello della registrazione, lieve stridore] / Come nuovo [minime imperfezioni del supporto fisico che non influiscono in alcun modo sulla riproduzione dell’audio]

6 [OTQ] – Questa definizione si applica a tutte le tipologie di supporti contenute nell’archivio? Indica qui se ci sono delle categorie che divergono sensibilmente dalla definizione generale, usando il formato [Tipologia: stato di conservazione, ad es. Cilindri: Molto buono].

7 [FC] – Sono state necessarie operazioni di restauro per i supporti fisici?

Sì <To Page 3> / No <To Page 4>

¹⁶ At the beginning of each page, a brief subhead explained to respondents how to use the ND option, i.e., “per le domande di cui non sai la risposta esatta, o per quelle che non trovano risposta date le tue scelte precedenti” and, in the case of personal information, “per quelle a cui semplicemente non vuoi rispondere”.

[P3] Consistenza dell'archivio fisico [Restauro]

- 8 [OTQ] – Che tipo di operazioni di restauro?
9 [OTQ] – Quando sono state effettuate le operazioni di restauro?
10 [FC] – È stato seguito un protocollo operativo documentato?
Sì / No
11 [FC] – Il protocollo è depositato insieme all'archivio?
Sì / No

[P4] Consistenza dell'archivio digitale

- 12 [FC] – Il tuo archivio è composto anche da documenti digitali (ad es, .mp3, .wav ecc.)?
Sì <To Page 5> / No <To Page 9>

[P5] Consistenza dell'archivio digitale [2]

- 13 [OTQ] – Quanti file audio sono contenuti nell'archivio?
14 [CATA] – In quale formato audio sono stati prodotti?
.mp3 / .aac / .ogg / .flac / .alac / .wav / Altro...
15 [OTQ] – Sai quantificare il numero di file per formato? Compila il riquadro qui sotto nel formato [Formato: numero file]
16 [FC] – Sono previsti incrementi?
Sì / No / ND
17 [OTQ] – In che qualità audio sono stati prodotti? Compila il riquadro qui sotto nel formato [Formato: bit depth/sampling rate]
18 [FC] – Alcuni di questi file audio sono digitalizzazioni di supporti analogici? (NB: i contenuti di CD/DVD/DAT sono considerati già documenti digitali)
Sì <To Page 6> / No <To Page 7>

[P6] Consistenza dell'archivio digitale [Digitalizzazioni]

- 19 [OTQ] – Quando sono state effettuate le digitalizzazioni degli originali?
20 [FC] – La digitalizzazione è avvenuta presso terzi?
Sì / No
21 [FC] – È stato seguito un protocollo operativo documentato?
Sì / No
22 [FC] – Il protocollo è depositato insieme all'archivio?
Sì / No

[P7] Consistenza dell'archivio digitale [Deposito online]

- 23 [FC] – Oltre o in sostituzione ad essere archiviati su supporti fisici, i file sono depositati in servizi di archiviazione online? (Ad es., servizi cloud come Drive)
Sì <To Page 8> / No < To Page 9>

[P8] Consistenza dell'archivio digitale [Deposito online 2]

24 [OTQ] – In quali servizi di archiviazione online è depositato l'archivio?

[P9] Altre generalità sull'archivio sonoro

25 [FC] – I documenti sonori sono stati in qualche modo catalogati?

Sì <To Page 10> / No <To Page 11>

[P10] Altre generalità sull'archivio sonoro [Ordinamento]

26 [CATA] – Quali criteri di catalogazione sono stati adottati?

Alfabetico / Cronologico / Numerico / Per autore / Tematico / Topografico / Altro...

[P11] Altre generalità sull'archivio sonoro [2]

27 [FC] – I singoli documenti contenuti nell'archivio sono stati descritti? (Ad es., data e luogo di raccolta, contenuto ecc.)

Sì <To Page 12> / No <To Page 13>

[P12] Altre generalità sull'archivio sonoro [Descrizione]

28 [OTQ] – Quali criteri di descrizione hai adottato per il tuo archivio (ad es., standard come ISAD(G), altro)?

[P13] Altre generalità sull'archivio sonoro [3]

29 [FC] – Sono stati attuati interventi per la conservazione a lungo termine dei materiali, sia per quanto riguarda i supporti (deterioramento fisico) che i documenti digitali (incompatibilità di formati)?

Sì <To Page 14> / No <To Page 15>

[P14] Altre generalità sull'archivio sonoro [Conservazione a lungo termine]

30 [OTQ] – Quali interventi per la conservazione a lungo termine sono stati attuati?

[P15] Altre generalità sull'archivio sonoro [4]

31 [FC] – Ogni quanto i materiali (o una loro parte) vengono ascoltati?

Una o più volte alla settimana / Una o più volte al mese / Una o più volte ogni sei mesi / Una o più volte all'anno / Mai o quasi mai

32 [FC] – A quando risale l'ultimo ascolto dei materiali (o di una loro parte)?

Da oggi a una settimana fa / Da una settimana fa a un mese fa / Da un mese fa a sei mesi fa / Da sei mesi fa a un anno fa / Più di un anno fa

[P16] Racconta l'archivio

33 [OTQ] – Quando sono state effettuate le registrazioni?

34 [OTQ] – In che località sono state effettuate le registrazioni?

35 [CATA] – Qual è l'area disciplinare di pertinenza dell'archivio?

Fonetica-fonologia / Dialettologia / Storia orale / Antropologia / Sociologia / Psicologia / Linguistica applicata / Sociolinguistica / Etnomusicologia / Scienze politiche / Musicologia / Altro...

36 [OTQ] – Quali lingue sono rappresentate nell'archivio? (Se si tratta di dialetti, specificare con la formula "Dialetto di [toponimo]")

37 [CATA] – Quali generi testuali sono rappresentati nell'archivio?

risposta a questionario linguistico (fonetico, morfologico, sintattico, lessicale) / risposta a questionario etnografico / risposta a questionario demologico (es. API) / lettura (di liste di parole, di fraseologia, di brano) / testimonianze (su un evento storico, sulla vita di fabbrica, su lotte politiche...) / storie di vita (narrazione della vicenda di una persona) / etnotesto / intervista / conversazioni spontanee tra più soggetti / dialoghi semi-spontanei (task di laboratorio, ad es. map-task) / descrizione di immagini e-o oggetti / canzone narrativa / canzone iterativa (canto enumerativo e iterativo) / canto itinerante di questua / canto lirico monostrofico / zingaresca / ninne nanne / filastrocche - cantilene / contrasti (in ottava) / scherzi - componimenti burleschi - canzoni di giochi / indovinelli / scioglilingua / detti e proverbi / invocazioni e scongiuri / poesia / poesia religiosa / grido / fiaba - favola - novella - racconto / leggende / rituali con formula [medicina popolare] / rito / teatro / teatro popolare / maggio drammatico (o epico) / Bruscello / Vecchia o Sega-la-Vecchia / Danza carnevalesca - ciclo carnevalesco / evento performativo di cultura popolare / discorso politico / conferenza / canzone pop / Altro...

38 [CATA] – Quali argomenti vengono trattati nei documenti sonori?

abitazione rurale / abbigliamento / agricoltura / alimentazione - cibi tradizionali / ambiente / anarchismo / aneddoti / animali / antifascismo - fascismo / archeologia industriale / arte / artigianato / attività dell'archivio stesso / attività didattiche / balie / befanata sacra / befanata profana / biografie / blasoni popolari / bonifica / bosco / brigantaggio / canti e musica popolari / canti di cantastorie / carbonai / carnevale / casa e attività domestiche / ciclo produttivo del castagno / cinema / coltellinaio / coltivazione / comizi / commemorazioni - anniversari - cerimonie / comportamento umano / condizione femminile / convegni - conferenze - dibattiti / corpo umano / cultura contadina / cultura di fabbrica / cultura materiale / danza popolare / dialetti e varietà di lingua / dopoguerra (secondo dopoguerra) / emigrazione / fabbro / famiglia mezzadrile / famiglia tradizionale / feste / feste popolari / feste religiose / feste e cerimonie tradizionali / filatura / focarazza / fonetica e fonologia dialettale / giochi e giocattoli per l'infanzia / gioco delle differenze - gioco delle mappe - gioco della telefonata - gioco dell'appuntamento / immigrazione / intercultura / intrecciatura / lavorazione del marmo / lavoro / lavoro contadino / lavoro nei boschi / lavoro - vita operaia / lessico dialettale / letteratura / letteratura popolare /

lezioni, educazione permanente, 150 ore / lotte (manifestazioni) sindacali e politiche / maggio lirico / maggio religioso / magia / manifestazioni / manifestazioni folcloriche / manifestazioni musicali / Maremma / medicina popolare / mestieri / mestieri tradizionali / meteorologia / mezzadria / migrazioni / militanza politica / miniera / miti e leggende / morfologia dialettale / mostre / movimento operaio - sindacale / museografia / musica etnica / nazismo - deportazioni - stragi - persecuzioni razziali / onomastica / passione di Cristo / pastorizia / pesca / piante e frutti / politica / prigionia / prima guerra mondiale (grande guerra) / prostituzione / razzismo / registrazioni televisive e radiofoniche / religione / religiosità popolare / resistenza (Liberazione) / ricordi di infanzia / riforma agraria / riti / seconda guerra mondiale / semantica dialettale / sindacalismo - militanza politica / sintassi dialettale / social network / società preindustriale / sport / storia delle donne / storia locale (rurale) / storia locale (urbana) / storia politica / storie cantate / stregoneria / teatro / teatro popolare / tecniche agricole tradizionali / tecniche di produzione industriale / tecniche di produzione pre-industriale / tecnologia / tempo (sua scansione) / territorio / tessitura / toponomastica / tossicodipendenza / tradizioni / tradizioni contadine / traduzione - interpretazione in dialetto di un brano / uomo (età della vita, relazioni sociali) / veglie / videogiochi / vita della comunità / vita quotidiana / Altro...

39 [OTQ] – Come e perché è nato l'archivio?

40 [OTQ] – Quali erano gli obiettivi della raccolta dei documenti sonori?

41 [OTQ] – Come sono stati reclutati i parlanti registrati?

42 [OTQ] – In quali contesti sono stati registrati i parlanti?

43 [FC] – La raccolta era finanziata?

Sì <To Page 17> / No <To Page 18>

[P17] Racconta l'archivio [Finanziamenti]

44 [OTQ] – Da chi era finanziata la raccolta?

[P18] Altri allegati

45 [FC] – Nell'archivio sono presenti anche documenti di altro genere connessi ai documenti sonori? (testo, foto, video; trascrizioni escluse)

Sì <To Page 19> / No <To Page 20>

[P19] Altri allegati [Testo, foto, video]

46 [CATA] – Di che tipo di documenti si tratta?

Testo / Video / Immagini - foto / Altro...

47 [OTQ] – Sai quantificare il numero di documenti per tipologia? Compila il riquadro qui sotto nel formato [Tipologia: numero documenti]

[P20] Altri allegati [2]

48 [FC] – I documenti sonori presenti nell'archivio sono stati trascritti e/o annotati?

Sì <To Page 21> / No <To Page 22>

[P21] Altri allegati [Trascrizioni e annotazioni]

49 [OTQ] – Chi ha trascritto e/o annotato i documenti?

50 [OTQ] – In percentuale sul totale, quanti documenti sono stati trascritti e/o annotati?

51 [CATA] – In che formato si presentano le trascrizioni e/o annotazioni?

Cartaceo / File .doc / File .txt / File .pdf / File .TextGrid (PRAAT) / File .eaf (ELAN) / File .tag (TranscriberAG) / Altro...

[P22] Questioni legali [1]

52 [FC] – L'archivio è stato oggetto di una dichiarazione di interesse storico particolarmente importante ai sensi del Codice dei beni culturali e del paesaggio? [HT]

Sì <To Page 23> / No <To Page 24>

[P23] Questioni legali [Dichiarazione d'interesse storico]

53 [OTQ] – In che anno è stato oggetto di dichiarazione di interesse storico particolarmente importante?

[P24] Questioni legali [2]

54 [FC] – Sui documenti dell'archivio ricadono dei diritti d'autore? [HT]

Sì <To Page 25> / No <To Page 26> / ND <To Page 26>

[P25] Questioni legali [diritti d'autore]

55 [OTQ] – Puoi dirci qualcosa sui soggetti titolari dei diritti d'autore? [HT]

[P26] – Questioni legali [3]

56 [FC] – Nei documenti dell'archivio sono contenuti dati personali? (“qualsiasi informazione riguardante una persona fisica identificata o identificabile, [...] si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale”) [HT]

Sì / No / ND

57 [FC] – Nei documenti dell'archivio sono contenuti dati supersensibili? (“dati personali che rivelino l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, nonché [...] dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona”) [HT]

Sì / No / ND

58 [FC] – Sono disponibili liberatorie firmate dai soggetti registrati?

Sì <To Page 27> / No <To Page 28>

[P27] – Questioni legali [Liberatorie]

59 [FC] – Le liberatorie sono depositate nell'archivio?

Sì / No

60 [CATA] – Quali sono gli aspetti regolamentati dalle liberatorie?

Privacy / Trattamento dei dati / Responsabili del trattamento dei dati / Finalità / Modalità di reclamo / Volontarietà della partecipazione / Compenso / Diritti d'autore - di esecuzione / Proprietà intellettuale / Modalità di riutilizzo / Altro...

[P28] Pubblicazioni

61 [OTQ] – I materiali dell'archivio sono stati poi utilizzati per pubblicazioni?

Sì <To Page 29> / No <To Page 30>

[P29] Pubblicazioni [2]

62 [CATA] – Di che tipo di pubblicazioni si tratta?

Articoli scientifici / Monografie / Articoli di giornale / Post sul web / Cataloghi / Corpus / Antologie / Altro...

[P30] Formazione pregressa [1]

63 [FC] – Prima di effettuare le registrazioni audio contenute nell'archivio, il raccoglitore aveva ricevuto una formazione tecnica (strumenti di registrazione, formati audio ecc.)?

Sì <To Page 31> / No <To Page 32> / ND <To Page 32>

[P31] Formazione pregressa [tecnica]

64 [CATA] – Che tipo di formazione tecnica aveva ricevuto il raccoglitore prima di effettuare le registrazioni audio contenute nell'archivio?

Corsi universitari / Corsi privati / Scuole estive / Studi autonomi / Altro...

[P32] Formazione pregressa [2]

65 [FC] – Prima di effettuare le registrazioni audio contenute nell'archivio, il raccoglitore aveva ricevuto una formazione archivistica?

Sì <To Page 33> / No <To Page 34> / ND <To Page 34>

[P33] Formazione pregressa [archivistica]

66 [CATA] – Che tipo di formazione archivistica aveva ricevuto il raccoglitore prima di effettuare le registrazioni audio contenute nell'archivio?

Corsi universitari / Corsi privati / Scuole estive / Studi autonomi / Altro...

[P34] Formazione pregressa [3]

67 [FC] – Prima di effettuare le registrazioni audio contenute nell'archivio, il raccogliitore aveva ricevuto una formazione legale?

Sì <To Page 35> / No <To Page 36> / ND <To Page 36>

[P35] Formazione pregressa [legale]

68 [CATA] – Che tipo di formazione legale aveva ricevuto il raccogliitore prima di effettuare le registrazioni audio contenute nell'archivio?

Corsi universitari / Corsi privati / Scuole estive / Studi autonomi / Altro...

[P36] Dichiarazioni d'interesse¹⁷

69 [FC] – Eri a conoscenza dell'esistenza del Vademecum?

Sì / No

70 [FC] – Saresti interessato a partecipare ad eventi di formazione relativi al mondo degli archivi orali?

Sì (avvertimi quando la mailing list del Tavolo permanente per le fonti orali sarà attiva) / No

71 [FC] – Saresti interessato a depositare il tuo archivio (digitale) in una piattaforma (repository) online?

Sì / No

72 [CATA] – Quali caratteristiche della piattaforma (repository) ti incentiverebbero a depositare il tuo archivio?

Interfaccia grafica intuitiva / Presenza di personale di supporto / Deposito gratuito / Indicizzazione dei documenti / Struttura archivistica predefinita secondo standard consolidati / Servizi web di restauro audio / Servizi web di annotazione audio / Versionamento dell'archivio / Livelli di accesso alle risorse personalizzabili / Possibilità di produrre rappresentazioni grafiche dei materiali (ad es., mappe di dati georeferenziati) / Collegamento con una comunità di ricerca affine / Altro...¹⁸

73 [FC] – Vuoi rimanere aggiornato sulla piattaforma *Archivio Vi.Vo.*?

Sì (avvertimi quando la mailing list di *Archivio Vi.Vo.* sarà attiva) / No¹⁹

¹⁷ This page had another subhead briefly describing the *Vademecum* (“Il Vademecum per il trattamento delle fonti orali - Un gruppo di lavoro interdisciplinare ha recentemente pubblicato un Vademecum [HT] contenente informazioni pratiche sulla produzione, conservazione e riuso delle fonti orali. Un Tavolo permanente [HT] si è in seguito costituito per mantenere aggiornato il testo e aggiungere progressivamente sezioni d'interesse per l'opera”).

¹⁸ After Item 72, in the same page, the *Archivio Vi.Vo.* platform was briefly introduced (“*Archivio Vi.Vo.* - *Archivio Vi.Vo.* è una piattaforma sviluppata nel corso di un recente progetto di Regione Toscana (2019-2021) e dedicata agli archivi orali. Grazie ad un'interfaccia grafica intuitiva, *Archivio Vi.Vo.* conduce l'utente attraverso i passaggi di un flusso di lavoro che porta dalla creazione di copie di conservazione a lungo termine alla descrizione di documenti d'accesso ordinati secondo consolidati standard archivistici. La piattaforma si avvale della rete europea CLARIN [HT] e sarà presto rilasciata pubblicamente”).

¹⁹ The following items were introduced through the following paragraph (“Per una rivista italiana degli archivi orali - In tempi recenti, il *Journal of Australian Linguistics and Language* si sono interrogati sulla possibilità di rendere il processo di archiviazione di documenti orali un prodotto della ricerca a sé stante. I fini ultimi sono incentivare la cura dei propri materiali e promuoverne l'accessibilità”).

74 [FC] – Saresti interessato alla formazione di una rivista italiana sulla costituzione di archivi orali?

Sì / No

75 [FC] – La possibilità di far valere il processo di costituzione di archivi orali come un prodotto della ricerca ti incentiverebbe a prenderti cura dei tuoi documenti sonori?

Sì / No

[P37] Anagrafica del referente²⁰

76 [OTQ] – Nome e cognome del referente

77 [OTQ] – Età

78 [OTQ] – Sesso

79 [FC] – Ruolo

Studente triennale / Studente magistrale / Dottorando / Borsista di ricerca / Assegnista - contrattista di ricerca / Ricercatore a tempo determinato / Professore associato / Professore ordinario / Ricercatore indipendente / Pensionato / Appassionato / Altro...

80 [OTQ] – Affiliazione accademica/istituzionale

81 [OTQ] – Appartenenza ad associazioni di ricerca

82 [CATA] – Settore di ricerca

Fonetica - fonologia / Dialettologia / Storia orale / Antropologia / Sociologia / Psicologia / Linguistica applicata / Sociolinguistica / Etnomusicologia / Scienze politiche / Musicologia / Altro...

83 [OTQ] – Indirizzo

84 [OTQ] – Telefono

85 [OTQ] – Sito web

86 [FC] – Sei anche il proprietario dell'archivio?

Sì <To Page 38> / No <To Page 39>

[P38] Anagrafica del referente [Proprietà]

87 [FC] – Come sei diventato proprietario dell'archivio?

Propria produzione / Donazione / Prestito / Acquisto

<To Page 40>

[P39] Proprietario dell'archivio

88 [OTQ] – Puoi dirci qualcosa sul proprietario dell'archivio?

[P40] Anagrafica del referente [Produzione]

89 [FC] – Sei anche il produttore dell'archivio? (ovvero, colui che ha prodotto, accumulato e/o conservato e usato la documentazione archivistica nello svolgimento della propria attività

²⁰ The usual subhead about ND usage was prefaced by a brief definition of 'referente' as the person filling in the form.

personale o istituzionale) [HT]

Sì <To Page 42> / No <To Page 41>

[P41] Produttore dell'archivio

90 [OTQ] – Puoi dirci qualcosa sul produttore dell'archivio?

[P42] Anagrafica del referente [Raccolta]

91 [FC] – Sei anche il raccoglitore dei materiali che costituiscono l'archivio? (ovvero, colui che ha effettuato le registrazioni audio contenute nell'archivio)

Sì <To Page 44> / No <To Page 43>

[P43] Raccoglitore dei materiali contenuti nell'archivio

92 [OTQ] – Puoi dirci qualcosa sul raccoglitore dei materiali contenuti nell'archivio?

[P44] Anagrafica dell'archivio

93 [OTQ] – Luogo/luoghi di conservazione dell'archivio (indirizzo)

94 [FC] – Consultabilità

Totale <To Page 45> / Parziale <To Page 45> / Negata <To Page 46>

[P45] Anagrafica archivio [Consultabilità]

95 [CATA] – Modalità di consultazione

Accesso libero online / Accesso libero previa autenticazione / Condivisione online previo invio di progetto di ricerca e-o motivazione specifica / Consultazione libera in situ / Consultazione in situ previo appuntamento e motivazione / Altro...

<To Page 47>

[P46] Consultazione negata

96 [CATA] – Qual è il motivo dell'impossibilità di consultazione?

Fragilità dei supporti / Indisponibilità strumenti di riproduzione / Indisponibilità risorse umane / Indisponibilità spazi / Tutela della riservatezza / Altro...

[P47] Anagrafica archivio [Conservazione]

97 [FC] – Chi è il responsabile della conservazione dell'archivio?

Nessuno / Il referente (io) / Il proprietario dell'archivio (se diverso dal referente) / Il produttore dell'archivio (se diverso dal referente e dal proprietario) / Un altro individuo / Un'università / Un istituto / Un repository / Altro <The first four options lead to Page 49, the others to Page 48>

[P48] Responsabile della conservazione

98 [OTQ] – Puoi dirci qualcosa sul responsabile della conservazione dell'archivio?

[P49] Anagrafica archivio [Questioni legali]

99 [FC] – Chi è il responsabile delle questioni legali connesse all'archivio?

Nessuno / Il referente (io) / Il proprietario dell'archivio (se diverso dal referente) / Il produttore dell'archivio (se diverso dal referente e dal proprietario) / Un altro individuo / Un'università / Un istituto / Un repository / Altro <The first four options end the form, the others lead to Page 50>

[P50] Responsabile delle questioni legali

100 [OTQ] – Puoi dirci qualcosa sul responsabile delle questioni legali relative all'archivio?