

10. Philology in the preservation of audio documents

Customized versus ready-made approaches

Federica Bressan

Sound recordings have proven to be irreplaceable primary sources for disciplines like linguistics, musicology, ethnomusicology and sociology. Their fragile physical nature has activated a number of counter-actions aimed at prolonging the life expectancy of their content. Methodological issues have been raised in the past three decades, considering the relationship between the physical object and its (digitized) intangible content, which is not only complex but develops over time. This article reflects on the role of the emerging discipline known as ‘digital philology’ in the long-term preservation of audio documents, pointing out how some concepts (such as *authenticity, reliability and accuracy*) may require a ‘customized’ (as opposed to a ‘ready-made’) approach in the preservation workflow – mainly depending on the type of the archive: unique copies, field recordings, electronic music, oral history, to name some representative cases.

The set-up of the laboratory for sound preservation at the Centro di Sonologia Computazionale (CSC) of the University of Padova, Italy, represents one customized approach in which conscious methodological decisions support philologically informed digitization efforts. The methods affect the results, and ultimately the consequences are not merely technological but cultural.

How inscriptions are photographed and text corpora are transcribed and encoded, as well as how a sound recording is re-mediated,¹ are “crucial for the way in which these research objects will be studied in the future” (Van Peursen, 2010, p. 11). In other words, the digital representation of data (the bit stream) and the organization and presentation of data (the cultural interfaces [Manovich, 2001]) are not neutral with respect to the final perception that users have of the ‘real’ or ‘original’ object, which may often not be available for comparison. In addition, the relationship between the digital objects and the real objects they allegedly represent “is not only very complex, it also develops over time” (Van Peursen, 2010, p. 10). In order to ensure that the electronic sources – which scholars and the general public are getting more and more accustomed to referring to for their research and personal interests – meet the requirements of authoritativeness, accuracy and reliability, it is necessary to define what makes a digital document authoritative, accurate and reliable. This definition implies an understanding of what the document means in its cultural context, and it is therefore not limited to technical questions that only computer scientists and engineers should be in charge of. ‘Digital philology’ is the multidisciplinary discipline that addresses these problems. While philology (“without

¹ ‘Re-mediation’ is the process of transferring acoustic information from one medium to another.

adjectives”, as Leonardi [2007, p. 65] puts it) has a long tradition, the reflections about digital philology are very recent, and so is the academic production around it. There is much work to be done in order to provide new methodological and operative tools to scholars whose aim is to author electronic editions based on digital sources. The shift introduced by the electronic medium regarding how texts are coded and accessed goes beyond their presentation, affecting the perception of the content.

The existing scientific literature about digital philology is mainly limited to bibliographic sources (with very few exceptions, such as Zattra, 2006, 2007). In this article, the term ‘text’ is sometimes extended to include audio recordings. In fact, one of the aims of philology is to reconstruct an original text based on variant copies of manuscripts: making a parallel in the audio domain, any recording of the same event can be considered a variant copy of the ‘original’ event, thus being the equivalent of a manuscript for textual criticism.

Unlike the fields of text encoding, analysis and philology, where a longer tradition provides a comprehensive research literature, long-term preservation of audio documents is relatively new and lacks a background of knowledge and experience. Audio documents have gained the status of documentary sources only recently, and it is not uncommon to find that texts are still considered ‘first-class’ sources while sound recordings are considered ‘second-class’ sources for scholarly studies. Unless certain criteria ensure that digital audio (and other multimedia) can be considered authoritative, accurate and reliable sources, this trend may be hard to get rid of in the future.

Besides the ongoing multiplication of digitized documents, the number of born-digital documents is also increasing, making the definition of procedures for storage and cataloguing even more urgent because a ‘physical original’ is never available for comparison.

This article presents some reflections on the problem of digital philology applied to audio documents in long-term preservation.

Preparation of sources

Even an image capture and editing, which may at first sight be a rather straightforward and ‘objective’ procedure... require[s] intellectual, critical choices, interpretation, and manipulation.

Dahlström (2010)

In order to plan and to perform trans-coding, it is necessary to have a model of the object or document. Computer science and philology may collide over how to define that model, because “humanities generally show terminological ambiguity

due to the heterogeneous and elusive object of study”, while computer science deals by definition with the “processing of data (encoded information) expressed in non-ambiguous languages” (Gigliozzi, 2003, p. 48). Creating a model of the object to be digitized means analysing it and selecting what relevant features will represent it. It is important to keep in mind that the creation of the model is not required exclusively for the sake of the computer: it should rather be seen as an “important space for analysis and for the formalization of the knowledge about the subject of the study” (p. 53).

The definition of the model is a ‘crucial’ part of the preliminary activities (see Van Peursen, 2010, p. 11): any further analysis will manipulate the electronic representations of the original physical objects. Gigliozzi (2003) suggests that a model already underlies any written text (a code for graphic symbols, syntax, narrativity, language, etc.) and that it can be useful to reflect on what features make it effectively represent the message (information) intended by the author. Starting the transition to the audio domain, the notion of ‘text’ is among the first to be defined. Is a recording of an acoustic event or electronic composition an expression or witness of the archetypical sound? And is there such a thing as an archetypical sound? Every audio document may be considered a ‘master’ recording in the sense that it bears witness to the acoustic event in a unique way, due to the manipulation and conditions that the document has been subject to through the years, including the aging of the carrier (see Bressan and Canazza, 2013; IASA Technical Committee, 2004). Without attempting to say that an archetypical sound exists, each recording may be seen as the philological equivalent of the written ‘text’.

Considering that the equivalent of textual criticism is performed by musicologists, linguists and most often experts other than computer scientists, the preservation task finishes when the equivalent of the diplomatic editions are ready. The process of preservation goes from the diagnosis of the physical document to the preparation of the digitized document for access (resources and services). The internal organization of the elements constituting the document may be modified during the cataloguing, in function of the contents: the type of object that is produced directly from the digitization is referred to as a ‘preservation copy’ (IASA Technical Committee, 2004) and is by definition the equivalent of the diplomatic edition (see Bressan *et al.*, 2013b). According to the definition of diplomatic edition, the eventual mistakes or imperfections of the physical document are maintained in the digital copy because they provide information about the author’s creative process (e.g. erased words), about the history of transmission of the document, and about the aging of the document (from dog-eared pages to patches of mould). The implication is that no restoration is

allowed in the audio of the preservation copy (such as noise removal or speech enhancement). The details of the model always have to be documented and publicly accessible. Anyone who accesses the digital resources has the right to access this information. Resources that “do not declare their objective and their limits, nor the procedures employed, nor the quality of the data to which the procedures have been applied, are low-profile products” (Gigliozzi, 2003, p. 120) According to the same source, a low-profile product is one that “uses the potential of computer technology without reflection” (p. 121).

A valid critical edition must be based on diplomatic editions (Gigliozzi, 2003, p. 122) that meet the requirements of authoritativeness, accuracy and reliability. These three qualities represent

a primary concern in long-term preservation... [With physical documents,] trustworthiness was all wrapped up in the concept of authenticity so that an authentic document was also reliable and accurate. This is no longer true (Duranti, 2012).

Authenticity needs to be redefined for electronic documents because they cannot “be preserved as... unchanged resources: we have only the ability to reproduce them” (CASPAR Consortium, 2008), and (un)intentional modifications introduced at some point of the file manipulation may be very difficult to remove. “Authenticity cannot be recognized as given, once and for ever, within a digital environment”; it can only be “approached asymptotically” (Factor *et al.*, 2009).

According to InterPARES 3, authenticity refers to

the trustworthiness of a record that is what it purports to be, untampered with and uncorrupted: it must be based on its identity and integrity, and on the reliability of the records system in which it resides.

Reliability is the trustworthiness of a record as a statement of fact: it must be based on the competence of its author, its completeness, and the controls on its creation. Accuracy is the correctness and precision of a record’s content: it must be based on the above, and on the controls on content recording and transmission (Duranti, 2012).

Textual criticism and digital authenticity

What could be the digital equivalent of text comparison in the audio field? Traditionally, musicologists are trained in the

study of musical scores. Only a few who specialized in the twentieth-century repertoire consider audio recordings a relevant documentary source for their studies – although the score (when existing) has always a powerful gravitational attraction. How can recordings on different tapes be effectively compared? Is it easier to do so with digital files? What are the sound parameters that are relevant and meaningful to a musicologist? Are there any? Answering these questions is the only way to enable the development of truly innovative methods and tools to assist or automate aspects of the musicologist’s work.

One of the means to compare texts has always been to present them in parallel columns. For example, Origen’s third-century *Hexapla* presented six versions of the Old Testament in parallel alignment (Fig. 1a). By analogy, we can imagine a possible technique for comparing a single audio feature extracted from three different audio files of the same recording (Fig. 1b).³ How much training would musicologists require to become familiar with the common ways of displaying audio signals on a computer, ranging from spectrograms to the example in Fig. 1b, and taking advantage of them to advance knowledge about music? Very likely, it would make sense to compare not two audio files but rather two complex objects, such as two preservation copies (including the metadata and the accompanying files, such as pictures of the cover, etc.).

“In postulating a typology of Electronic Philology, we must take into account the data, the procedures, and the results”, writes Marcos Marín (2001, p. 16). Any computational analysis of texts is, by definition, quantitative (e.g. word count); therefore the lowest extent in which the computer can serve the philologist is by providing her with a great amount of data and by extracting other data from them. Any list of words, sorted in any order, can be a good example. Features such as the file duration or the signal’s average amplitude could be the audio equivalent.

These first, basic, useful results constitute a set from which secondary results can be obtained. In particular, secondary results can be selected from primary ones by using complex information retrieval systems and rich query languages that have been developed to exploit huge textual resources. An example is the list of selected words obtained from a whole list of words. Selecting all files with a maximum peak over -3 dB could be the audio equivalent. Finally, tertiary results are obtained from the selected (secondary), following an exact pattern. Human interpretation is crucial. Examples for texts include using a concordance or an index to build a

² On the International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 3 Project, see www.interpares.org/ip3/ip3_index.cfm.

³ The example was generated with Matlab MIR toolbox.

Hebrew.	Hebrew Transliterated.	Aquila.	Symmachus.	LXX.	Theodotion.	Variante.
קַרְנֵי יַגְלִיל	λαμανασα	τῶ νικοποιῶ	ἐπινίκιος	εἰς τὸ τέλος	τῶ νικοποιῶ	εἰς τὸ τέλος
בְּתֵי קֹרֶן	βηη κορε	τῶν υἱῶν κορέ	τῶν υἱῶν κορέ	ὑπὲρ τῶν υἱῶν κορέ (τοῖς υἱοῖς)	τοῖς υἱοῖς κορέ	
וְיֵשׁוּב	אל· αλαμωθ	ἐπὶ νεανιοτήτων	ὑπὲρ τῶν αἰωνίων	ὑπὲρ τῶν κρυφίων	ὑπὲρ τῶν κρυφίων	ψαλμός
וְיֵשׁוּב	σιρ	ἔσμα	ψῆθ·	ψαλμός	ψῆθ	
וְיֵשׁוּב	ελαεμ λαου	<δ θεος ἡμῶν>	ὁ θεος ἡμῶν	ὁ θεος ἡμῶν	ὁ θεος ἡμῶν	
וְיֵשׁוּב	μασα· ουος	ἐλπίς καὶ κράτος	πεποιθήσεις καὶ ἰσχὺς	καταφυγὴ καὶ δύναμις	καταφυγὴ καὶ δύναμις	
וְיֵשׁוּב	εθ	βοήθεια	βοήθεια	βοήθεια	βοήθεια	
וְיֵשׁוּב	βαραωθ	ἐν θλίψεσιν	ἐν θλίψεσιν	ἐν θλίψεσι	ἐν θλίψεσιν	
וְיֵשׁוּב	νεμα· μωθ	ἐνθεοῖς σφάρα	ἐνθεοῖς σφάρα	ταῖς ἐνθεοῖς ἡμάς (ἐνθεοῖς ἡμῶν)	ἐνθεοῖς σφάρα (ταῖς ἐνθεοῖς ἡμάς)	
וְיֵשׁוּב	אל· χεν	ἐπὶ τούτῃ	διὰ τοῦτο	διὰ τοῦτο	διὰ τοῦτο	
וְיֵשׁוּב	לח· נירא	οὐ φοβηθήσάμεθα	οὐ φοβηθήσάμεθα	οὐ φοβηθήσάμεθα	οὐ φοβηθήσάμεθα	
וְיֵשׁוּב	βααμיר	ἐν τῷ ἀνταλλάσσει· θαι	ἐν τῷ συγχέεισθαι	ἐν τῷ ταράσσεισθαι	ἐν τῷ ταράσεισθαι	
וְיֵשׁוּב	[a]lars	γῆν	γῆν	τὴν γῆν	τὴν γῆν	
וְיֵשׁוּב	ου βαμωτ	καὶ ἐν τῷ σφάλλεισθαι	καὶ κλίναςθαι	καὶ μετατίθεσθαι	καὶ σαλεύεισθαι (μετατίθεσθαι)	
וְיֵשׁוּב	αριμ	ἔρη	ἔρη	ἔρη	ἔρη	
וְיֵשׁוּב	βλεβ	ἐν καρδίῃ	ἐν καρδίῃ	ἐν καρδίῃ	ἐν καρδίῃ	
וְיֵשׁוּב	ιαμιμ	θαλασσῶν	θαλασσῶν	θαλασσῶν	θαλασσῶν	

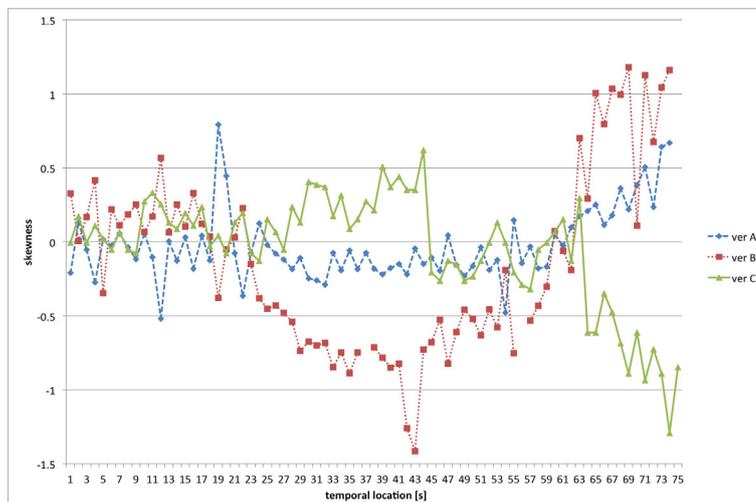


FIGURE 1a–b. Traditional presentation of the text in parallel columns (top) and possible presentation of multiple audio ‘texts’ (bottom) for comparison.

dictionary or the results of the collatio (the assemblage or collation of source texts) to prepare a critical edition (Marcos Marín, 2001, p. 16).

Towards a philologically informed methodology for preservation

Sound recordings have proven to be irreplaceable primary sources for disciplines like linguistics, musicology, ethnomusicology and sociology. The Centro di Sonologia Computazionale (CSC) in Padova has been active in the field of audio preservation and restoration for nearly two decades, building on a strong scientific background in sound synthesis and electronic music since the 1970s. The CSC houses a laboratory with the equipment required to create digital preservation copies (see Bressan and Canazza, 2013; IASA Technical Committee, 2004) that meet the requirements of accuracy, reliability and authenticity presented above. A protocol for re-mediation (Fig. 2) and a controlled environment are the key to quality control along the

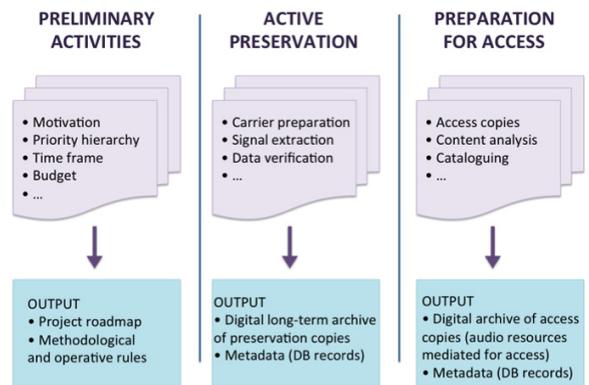


FIGURE 2. The scheme summarizes the main steps involved in the process of preservation of audio documents according to the methodology adopted at the CSC.

workflow (for more details on the protocol, see Bressan et al., 2013a). This goal is also achieved thanks to open-source software developed at the CSC (Bressan and Canazza, 2013), as well as with the multidisciplinary approach that

distinguishes the centre's methodology for audio preservation (involving information engineering, music and musicology, and chemistry). As Van Peursen (2010, p. 11) observes: "The creation of digital objects has to meet the standards of the various disciplines involved, and... is a crucial part of humanities research. It is more than just preparation for research".

The CSC laboratory features two working stations equipped with Apple desktop machines. The main station is dedicated to the analogue-to-digital (A/D) transfers and uses an A/D-D/A converter (Prism ADA-8XR) that supports 96 kHz/24-bit audio quality. Audio and video are acquired on separate machines (Fig. 3), and the entire re-mediation system is connected for safety to uninterruptible power supplies (UPS). A professional STUDER A810 with stereo heads is used to read most quarter-inch tapes. Digitized audio is exported in a non-compressed open format and stored in three different locations, after being analysed, and after the metadata have been automatically ingested into the database by a software especially developed at the CSC (Bressan and Canazza, 2013). The same software completes the preservation copy by processing the remaining data (checksums, images, video, etc.). The cataloguing staff receives automatic mail notifications whenever new material is available in the shared repository.

The laboratory also features a photographic set-up for the production of contextual information (front/back views as well as close-ups of all relevant details). The set-up was designed for short and frequent photo sessions, maximizing the quality of the picture with the minimum effort (1) to adjust the positioning of the camera and its parameters for each session; and (2) to transfer the new files to the desktop work station without dismantling the set-up or changing its configuration. The functionality of the set-up was based on the requirements reported by the Istituto Centrale per il Catalogo e la Documentazione (ICCD) and by the Italian Ministry of Culture in Galasso and Giffi (1998).

A precision incubator (Memmert INP 400) is used for the physical recovery of magnetic tapes. Thermal treatment, performed with the incubator, consists in applying consistent heat to the tapes over a specified period of time; it is aimed at reverting the effects of soft binder syndrome-sticky shed syndrome (SBS-SSS) (see Hess, 2008). Among the symptoms of this condition are sticking, squealing and abnormal shedding of the magnetic coating. It should be noted that not all tapes are suitable for thermal treatment (Bressan *et al.*, 2015). In order to find alternative recovery methods, chemical analyses and experiments are currently being performed by the CSC in collaboration with the Chemical Sector of the Department of Industrial Engineering at the University of Padova (Bressan *et al.*, 2015).

Conclusions

The intention of these reflections on the role of digital philology in the long-term preservation of audio documents is to encourage more discussions around the topic. Much research conducted in the field of computer science is preparatory to other scientific fields, such as (ethno-) musicology, linguistics and anthropology, mainly with methods and tools (e.g. information retrieval, signal processing, data compression). In the field of musical cultural heritage and digital libraries, computer science is instrumental in performing the re-mediation of audio documents and in managing digital data produced during re-mediation. Philology can contribute to the creation of audio digital libraries with some of its traditional concepts (e.g. authorship) and activities (e.g. textual criticism). 'Digital' philology has its own methods and meanings in the context of historical audio documents and archives, and its own considerations at the theoretical level. Unless these questions are addressed by the experts, the tendency to consider sound recordings 'second-class' documentary sources will continue. Applying the concepts of digital philology to electronic documents is vital to the creation of reliable digital libraries for text, audio, images, etc. In the audio field, the concept of text needs to be transferred to the preservation copy (not only the audio data but also contextual information and exhaustive metadata), and innovative methods and tools for text



FIGURE 3. Test for video shooting. On the left, a correct setting of the camera: each frame is clear and the text is readable. On the right, a wrong setting of the camera. The aim is to detect physical defects of the tape (due to aging or joints) that may correspond to a specific noise in the audio signal. Audio is directly recorded from a secondary output of the STUDER A810.

comparison need to be developed. Any analysis or study based on a preservation audio archive that does not ensure its accuracy, authoritativeness and reliability may lead to questionable results, critical scorn and even conclusions that may be potentially harmful for subsequent studies.

References

- Bressan, F., Bertani, R., Furlan, C., Simionato, F. & Canazza, S. 2016. An ATR FTIR and ESEM study on magnetic tapes for the assessment of the degradation of historical audio recordings. *Journal of Cultural Heritage*, 18: 313–320.
- Bressan, F. & Canazza, S. 2013. A systemic approach to the preservation of audio documents: methodology and software tools. *Journal of Electrical and Computer Engineering*, 2013 (article ID 489515): 21.
- Bressan, F., Canazza, S. & Bertani, R. 2015. Honey, I burnt the tapes! A study on thermal treatment for the recovery of magnetic tapes affected by Sticky Shed Syndrome. *IASA Journal*, 44: 53–64.
- Bressan, F., Canazza, S., Bertani, R., Roda, A. & Fontana, F. 2013a. The safeguard of audio collections: a computer science based approach to quality control in the archive of the Arena di Verona. *Advances in Multimedia*, 2013 (article ID 276354): 14.
- Bressan, F., Canazza, S., Roda, A., Bertani, R. & Fontana, F. 2013b. Pavarotti sings again: a multidisciplinary approach to active preservation of the audio collection at the Arena di Verona. *Journal of New Music Research*, 42(4): 364–380.
- CASPAR Consortium. 2008. Report on OAIS access model. Technical report, Centre national de la recherche scientifique and Université de Technologie de Compiègne.
- Dahlström, M. 2010. Critical editing and critical digitisation. In W.T. van Peursen, E.D. Thoutenhoofd & A. van der Weel, eds. *Text comparison and digital creativity: the production of presence and meaning in digital text scholarship*, pp. 79–98. Leiden, Brill.
- De Stefano, P., Tarr, K., Buchman, M., Oleksik, P., Moscoso, A. & Moskowitz, B. 2013. Digitizing video for long-term preservation: An RFP guide and template. Technical report, Barbara Goldsmith Preservation and Conservation Department, New York.
- Duranti, L. 2012. Interpares3 – Team Canada final report. Technical report, University of British Columbia.
- Factor, M., Henis, E., Naor, D., Rabinovici-Cohen, S., Reshef, P., Ronen, S., Michetti, G. & Guercio, M. 2009. Authenticity and provenance in long term digital preservation: modeling and implementation in preservation aware storage. Paper presented at the First Workshop on the Theory and Practice of Provenance, San Francisco, 23 February 2009.
- Galasso, R. & Giffi, E. 1998. La documentazione fotografica delle schede di catalogo – metodologie e tecniche di ripresa. Technical report, Istituto Centrale per il Catalogo e la Documentazione (ICCD) – Ministero per i Beni e le Attività Culturali, Rome.
- Gigliozzi, G. 2003. *Introduzione all'uso del computer negli studi letterari*. Milan, Bruno Mondadori.
- Hess, R. 2008. Tape degradation factors and challenges in predicting tape life. *ARSC Journal*, 39(2): 240–274.
- IASA Technical Committee. 2004. *Guidelines on the production and preservation of digital audio objects*. 1st ed. K. Bradley, ed. Standards, Recommended Practices and Strategies, IASA-TC 04. Aarhus, Denmark, International Association of Sound and Audiovisual Archives.
- Leonardi, L. 2007. Filologia elettronica tra conservazione e ricostruzione. In A. Ciula, ed. *Digital philology and medieval texts: [proceedings of the Arezzo Seminar 2006, 19–21 January]*, pp. 65–75. Arti spazi scritte 4. Ospedaletto (Pisa), Italy, Pacini Editore.
- Manovich, L. 2001. *The language of new media*. Cambridge, Massachusetts, USA, MIT Press.
- Marcos Marín, F. A. 2001. Where is electronic philology going? The present and future of a discipline. In D. Fiormonte & J. Usher, eds. *New media and the humanities: research and applications: proceedings of the first seminar on Computers, Literature and Philology Edinburgh, 7–9 September 1998*, pp. 11–22. Oxford, University of Oxford Humanities Computing Unit.
- Van Peursen, W. 2010. Text comparison and digital creativity: an introduction. In W.T. van Peursen, E.D. Thoutenhoofd & A. van der Weel, eds. *Text comparison and digital creativity: the production of presence and meaning in digital text scholarship*, pp. 1–28. Leiden, Brill.
- Zattra, L. 2006. The critical editing of computer music. In *Proceedings of the Electroacoustic Music Studies (EMS) conference*, p. 8. Beijing, Electroacoustic Music Studies Network.
- Zattra, L. 2007. The assembling of stria by John Chowning: A philological investigation. *Computer Music Journal*, 31(3): 38–64.

Federica Bressan is a post-doctoral researcher at Ghent University, where she leads a research project on interactive installation art under the Marie Curie funding programme H2020-MSCA-IF-2015. She holds an MA in musicology and a PhD in computer science. She previously held a post-doctoral research position at the Department of Information Engineering, University of Padua, Italy,

where she coordinated the laboratory for sound preservation and restoration. The core of her research lies in the study of the relationship between historical audio recordings and their digital representation – a relationship which is not only very complex but also develops over time (digital philology, hermeneutic

implications of encoding cultural objects). She is a member of the Steering Committee of the Ghent Centre for Digital Humanities (GCDH) and serves as General Chair of the International Workshop on Digital Philology for the Preservation of Multimedia Archives.