

Two Projects and a Thesaurus. Recent Experiences in the Management, Description and Indexing of Oral Sources

Sabina Magrini^(a)

a) Ministero della Cultura

Contact: Sabina Magrini, <u>sabina.magrini@beniculturali.it</u>
Received: 13 April 2021; Accepted: 15 July 2021; First Published: 15 January 2022

ABSTRACT

The Istituto Centrale per i Beni Sonori e Audiovisivi (ICBSA) has just finished, together with the Università degli Studi di Siena and Università degli Studi di Siena per stranieri, to work on the project "Ti racconto in italiano" which focuses on providing different access points to audio resources collected between 1980's and 2000's by ICBSA itself, as part of its mission to document Italian audio and audiovisual culture.

The main aim of the project is to create tools which will enable scholars of social history, art and literature to use these sources as well as providing original material for foreign students to exercise their knowledge of Italian. In order to facilitate access it has been necessary to create finding aids such as indexes and thesauri. For this purpose ICBSA has started a collaboration with the Biblioteca Nazionale Centrale di Firenze and the latter's Nuovo Soggettario.

This is not the first case of a project by institutes of the Italian Ministry of Culture comprising the use of the Nuovo Soggettario for the indexing of archival materials. Indeed, the Soprintendenza Archivistica e Bibliografica della Toscana has already worked in this direction a few years ago when treating the so-called Straw archives.

KEYWORDS

Nuovo Soggettario; Oral sources; Archives; Indexing.

^{© 2022,} The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. JLIS.it is a journal of the SAGAS Department, University of Florence, Italy, published by EUM, Edizioni Università di Macerata, Italy, and FUP, Firenze University Press, Italy.



The aim of this paper is to address a number of significant issues concerning the main theme of this Conference on bibliographic control in the digital ecosystem and mainly:

- 1. The complex interactions that are becoming common between different areas of knowledge and knowledge management in the bibliographic universe;
- 2. New ways of indexing documents;
- 3. The role of Thesauri in digital systems.

To do so, it shall be necessary to concentrate at first on the project "Archivi di paglia" which the Soprintendenza archivistica e bibliografica della Toscana developed around 2014-2016 in collaboration with the Biblioteca Nazionale Centrale of Florence. For those who may not be familiar with the intricacies of Italian cultural administration, the Soprintendenza is a Supervision Agency, the local office of the Italian Ministry for Culture engaged in the protection and valorisation of notified archives and libraries belonging to private individuals or archives and libraries belonging to public (non-State) entities in Tuscany. Object of this project was the census of the companies (and their archives) which produced straw hats in Tuscany in the past or still have connections to that world somehow.

Following, the project "Ti racconto in italiano" shall be illustrated. This is the result of the collaboration between the Istituto Centrale per i Beni Sonori e Audiovisivi (ICBSA, another office of the Ministry which concentrates its activity on the preservation and valorisation of audio and audiovisual heritage), l'Università di Siena (UNISI), l'Università di Siena per Stranieri (UNISTRA-SI) as well as the Biblioteca Nazionale Centrale of Florence (BNCF). This project is particularly interesting in this context as it bears agreat focus on indexing issues and, as it dates back to 2020, it adopts state of the art digital solutions.

These two projects have been chosen as interesting sample cases as they are both recent and quite unique in their sort. Cross-referencing between the worlds of library and archive databases is still relatively uncommon.

As concerns the "Archivi di paglia" project, it is necessary to illustrate at first the context in which the idea of such a research developed.

Archival records have a permanent significance for history, science and culture, as well as for the legal protection of individuals and legal entities. As such they can truly be considered to be a cultural asset.

Amongst the archival records that the Soprintendenza archivistica e bibliografica della Toscana safeguards there are, since the 1970's, business or company archives. Such archives bear witness to the history, capacities and vision of many big and small intrapreneurs in the local manufacturing industry. Tuscany has been famous since the 18th century for the production and processing of straw. The latter was used to realize the famous 'cappello di paglia di Firenze' (viz. the Florentine straw hat), giving work to hundreds of women employed in weaving straw into braids and hats. The tradition has continued somehow until today and some of the older firms are still active: in all, there are around 14 firms, many of which around a century old, which are still producing braids, hats and hat moulds. Their work and tradition has inspired the Museum of Straw in Signa¹, one of the main centres of the production of straw hats. It was in Signa that, for the first time, in 1714 Domenico Michelacci had the idea of starting a new kind of straw crop in order to obtain a thread that was particularly suited for weaving.

-

¹ https://www.museopaglia.it/ Accessed June 2021



To celebrate the third centenary of the revolution in straw production and processing in the region after Michelacci's pioneering experiences, the Soprintendenza archivistica opted in 2014 to realize the census of the companies that were or are still active in the sector, to collect data on their archives and to organize and record a series of interviews with business owners, workers and furnishers active in this line (as well as their relatives) in order to have first-hand information on their way of life and work. The interviews (some on video as well) were intended for the Museum at Signa. One of the main aims of the project was also the publication online of the so-called SIU-SA descriptions of the archival records of the firms involved. SIUSA is an acronym for the Sistema Informativo Unificato per le Soprintendenze Archivistiche (Unified Information System for the Supervision Agencies). It intends to be the primary access node to non State archival documents, both public and private, which are not kept by StateArchives.

The system describes the *archival fonds* according to a multi-level description; the *creators* (*bodies*, *people* and *families*) who produced the documents performing their activities; the persons or bodies who preserve (custody) the archives. General historical, administrative and archival information is provided as well, in order to allow a better comprehension of the context².

The aim of SIUSA is to assure the preservation and the knowledge of these sources and to provide access to them.

It was immediately clear from the perusal of the documents and the examination of the content of the interviews that it would have been essential to dispose of a controlled vocabulary focused on the world of straw processing. In such a way, it would have been possible to choose from a selection of terms in order to index content or to retrieve content through browsing or searching: thus, the SIUSA descriptions and any other works on the archives and interviews would have gained so much in sense, purpose and usability!

The potential of language as a meeting point between libraries, archives and museums was on the other hand a theme of reflection in those years for the Soprintendenza. So much so that in 2012 it had already created the basis of a collaboration with the Biblioteca Nazionale Centrale di Firenze through series of explorative letters and reciprocal declarations of intent.

For this reason, the Soprintendenza archivistica sought the collaboration of the Biblioteca Nazionale Centrale of Florence and specifically of the team behind the Nuovo soggettario³. The Nuovo soggettario viz. the New Subject Index, is the Italian subject indexing tool created by the National Central Library of Florence for the entire system of Italian libraries and, in particular, for those operating in the National Library Service.

Far from being a tool in use only in the Library world, the Nuovo soggettario was and still is open to contributions from other areas of knowledge management and is interoperable with databases of archives and museums as well as available in all standard formats and protocols.

Significantly the potential of this interaction between different worlds was explored on occasion of the conference organized in 2015 by ANAI, the Association of Italian Archivists⁴, MAB *Musei Archivi Biblioteche. Professionisti del Patrimonio Culturale Toscana* as well as the Tuscan Region and hosted by the Soprintendenza itself. On that occasion Emilio Capannelli, one of the archivists

² https://siusa.archivi.beniculturali.it/cgi-bin/siusa/pagina.pl?RicLin=en Accessed June 2021

³ https://thes.bncf.firenze.sbn.it/ Accessed June 2021

⁴ http://www.anai.org/anai-cms/; http://www.mab-italia.org/ Accessed June 2021



of the Soprintendenza, described the first experiences of collaboration between local archivists and librarians⁵.

So, thanks to this collaboration with the National Library, Alessia Artini and Silvia Melloni, the two free lance archivists who worked on the straw archives project under the guidance of the Soprintendenza (and of the archivists Renato Delfiol and Luca Faldi in particular) produced a series of controlled vocabulary terms which have been accepted and adopted by the general Thesaurus that is the main component of Nuovo soggettario system⁶. These terms hence recur in the scientific production concerning the straw archives.

The idea was that of continuing the collaboration between archivists and librarians in order to create other carefully selected lists of words and phrases in order to tag units of information in other 'domains', such as the archival records of other manufacturing sectors. So far, though, this has not occurred, at least as concerns the Soprintendenza archivistica in Tuscany.

A more complex and recent project that develops significantly some of the themes touched by "Archivi di paglia" is that of "Ti racconto in italiano" (fig. 1). This one year long project which has been financed by the General Directorate of Libraries in the Ministry aims at promoting the collections of ICBSA as well as the knowledge of the Italian language and culture abroad. What is it all about? And, first of all, what is ICBSA?

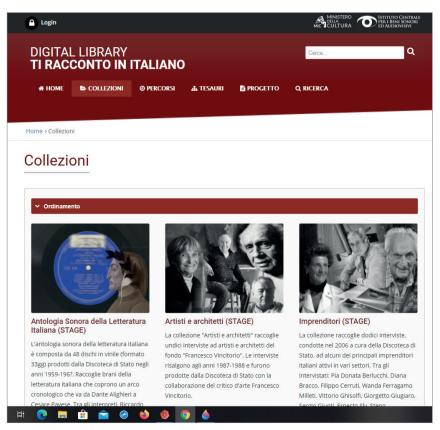


Fig. 1.

⁵ Capannelli 2016, 17-20

⁶ Artini, Benelli and Melloni 2017, 9



ICBSA – as stated before The Central Institute for Sound and Audiovisual Heritage⁷ – was founded in 1928 and was once known as the State Discotheque. Its first collections consisted of a record collection entitled "The word of the Great", voices collected by Rodolfo De Angelis in the first half of the 1920s. Over the years, documents of folklore, music, history, theater, dance, cinema have been added to this initial nucleus, which represented the first Italian public sound heritage, recorded on the most different media, from the wax cylinders invented by Edison, to records, tapes, videos up to current digital media. ICBSA materials are public and available for consultation via OPAC (Online Public Access Catalogue) with the possibility of listening to the *incipit* of the digitized sound documents and consultation of the accompanying description.

At the beginning of 2020, a collaboration between ICBSA, and the University of Siena (UNISI) and the University for Foreigners of Siena (UNISTRASI) started in order to make a section of the "Historical Voices" collection available to the public.

The materials made available by ICBSA for the realization of the project are 35 interviews on audio files lasting an average of one hour each, carried out between 1983 and 2006. The interviews, chosen by Piero Cavallari (an ICBSA technician) involve prominent personalities from the world of Italian business, art and culture for a total of about 40 hours of recording.

In detail, the corpus of interviews comprises:

- 13 interviews with writers, intellectuals, actors, directors (1983-1989): Elio Filippo Accrocca, poet; Giorgio Bassani, writer; Attilio Bertolucci; poet; Giorgio Caproni, poet; Riccardo Cucciolla, actor; Margherita Guidacci, poet; Luciano Lucignani, film director; Luciano Luisi, poet, writer andreporter; Mario Luzi, poet; Ettore Paratore, Latin scholar; Guglielmo Petroni, writer; Luisa Spaziani, poet; Franca Valeri, actress. (fig. 2)
- 11 interviews with artists (1987-1988): Carlo Belli, intellectual interested in art, architecture, music, archeology, politics; Maria Lai, designer; Carlo Lorenzetti, sculptor in metal; Teodosio Magnoni, painter and sculptor; Elisa Montessori, painter; Alberto Sartoris, architect; Ruggero Savinio, painter and author; Toti Scialoja, painter and poet; Guido Strazza, experimental artist; Giuseppe Uncini, painter and sculptor in iron and cement; Renzo Vespignani, painter, illustrator, set designer and engraver.
- 11 interviews with entrepreneurs (2006): Pia Berlucchi, wine producer; Diana Bracco, health and diagnosys; Filippo Cerruti, tourism, maritime transport, events; Wanda Ferragamo, fashion; Vittorio Ghisolfi, plastic producer; Giorgetto Giugiaro, design; Sergio Giunti, book editor; Ernesto Illy, coffee producer; Steno Marcegaglia, steal producer; Loris Meliconi, house goods producer; Ottavio Missoni, sportsman and fashion designer.

-

⁷ http://www.icbsa.it/ Accessed June 2021



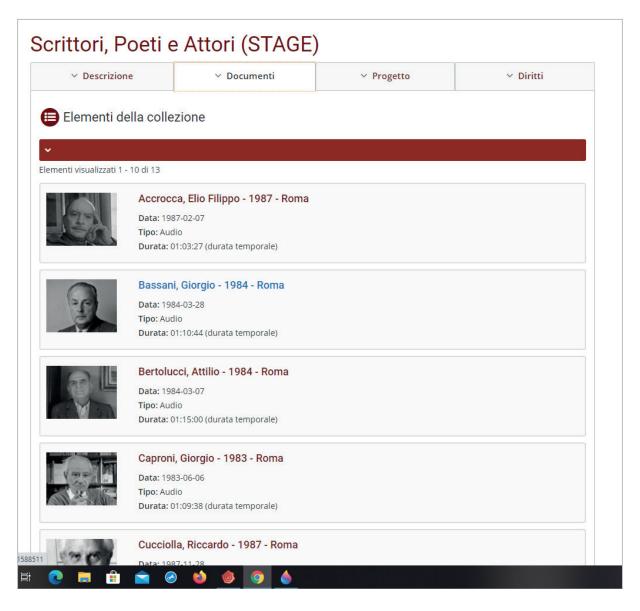


Fig. 2.

As concerns UNISTRASI, ICBSA decided to activate a research grant in order to carry out the didactic adaptation of these sound materials to make them available online to teachers and students of Italian, Level B1-B2. The purposes were two: to bring foreign learners closer to intensive listening to encourage the learning of Italian, but also to open them a world of ideas, stories and emotions linked to some extremely interesting Italian personalities of the second half of the 20th century and through them to offer a cross-section of the Italian culture and society of this period. The work was accomplished by Elena Grifoni under the supervision of Pierangela Diadori (Full Professor of Didactics of Italian for Foreigners).

As Pierangela Diadori has noticed, it is difficult to think that a person with a mother tongue other than Italian and unfamiliar with the Italian culture of the second half of the twentieth century



could listen to interviews lasting over an hour, or read the automatic transcription offered (honestly incomprehensible in many points): not even an Italian native speaker would do it.

The goal was therefore to create a battery of metadata, which could be freely surfed online and associated with the audio files and written files relating to each interview. It all had to be in the form of comprehension or completion exercises with closed answers, to be carried out in relation to the 'listening or reading' of the texts provided, together with keys for self-learning (figs. 3-4).

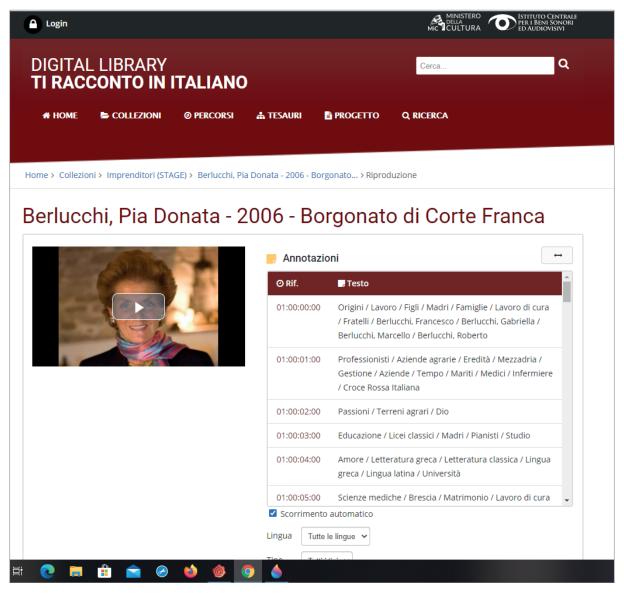


Fig. 3.



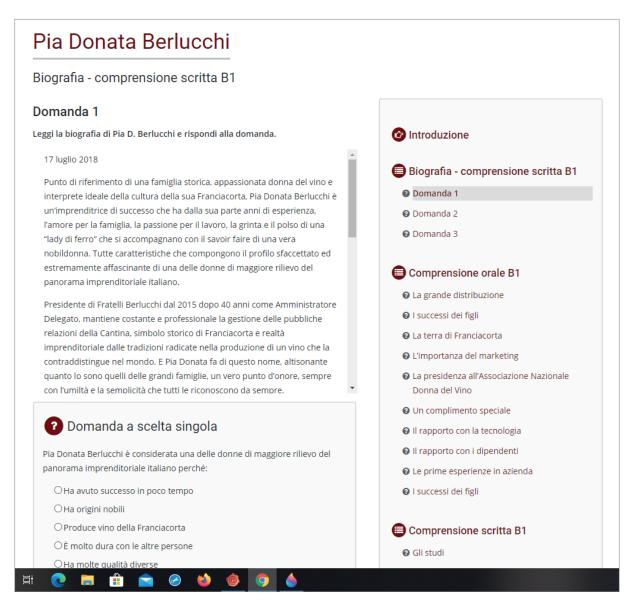


Fig. 4.

The format had to be such as to allow the insertion of these sets of exercises on the platform, in order to make access faster and offer immediate feedback with keys to the answers. Also a glossary has been realized.

If one thinks of the shortage of didactic audio materials, having this instrument available anywhere in the world online, with the advantage of being free of charge and the possibility of accessing authentic cultural contents, it is easy to understand the importance of this operation⁸.

UNISI's contribution to the project consisted, instead, in the indexing of the text in order to make it interesting for a wider number of potential users and perusable for research purposes. In this

⁸ Diadori 2021

JLIS.it vol. 13, no. 1 (January 2022) ISSN: 2038-1026 online Open access article licensed under CC-BY DOI: 10.4403/jlis.it-12731



case as well a research grant was activated: the work was accomplished by Cecilia Valentini under the supervision of Silvia Calamai (Associate Professor of Glottology and General Linguistics).

As Valentini and Calamai have pointed out on occasion of the presentation of the project at the CLARIN Annual Conference 2020⁹, the main objective of "Ti racconto in italiano" was to create tools facilitating users search through the collection. Therefore, finding aids such as indexes and thesauri have been realized.

Indexing has been done classifying each segment of the documents at regular time intervals with a label (fig. 3).

The terms used as labels consist of key words and controlled vocabulary and are structured in a specifically created thesaurus, viz. a specialized vocabulary of hierarchically listed words and phrases that indicates a preferred term among synonyms and shows relationships between terms. The use of a thesaurus facilitates retrieval of information and ensures greater consistency in the indexing of documents.

The thesaurus used is naturally that of the Nuovo soggettario¹⁰. The National Central Library of Florence is in fact a partner of the project. Therefore, Anna Lucarelli and her team have followed all phases concerning the choice of the terms and have authorized the use of new terms proposed by Cecilia Valentini and which have now entered the Nuovo soggettario.

Indexing has been done, mainly by Stella Montanari via AVIndexer, a software developed by Davide Merlitti (Informatica Umanistica, Pisa)¹¹ that makes use of SKOS, the Simple Knowledge Organization System recommended by W3C. Also the DublinCore Metadata set has been exploited (figs. 5-6).

Once indexed, the records as well as the thesauri will be published on the internet portal "Ti racconto in italiano" managed by ICBSA. The process is in progress at the moment. This platform is shaped on the digital library *Ti racconto la storia* online since September 2018¹². The latter was conceived by the General Directorate for Archives and the Central Institute for Archives in order to promote the knowledge and use of collections of oral testimonies, stories of life and other audio and audiovisual documentation produced on both analogue and digital media and stored in public institutions, research centers and private associations.

Once online, the perusal of the platform¹³ will clearly show how the resources of the Semantic Web have been exploited in order to create a framework for creating, managing, publishing and searching semantically rich metadata for web resources.

⁹ https://www.clarin.eu/content/programme-clarin-annual-conference-2020 Accessed June 2021

https://thes.bncf.firenze.sbn.it/ Accessed June 2021

¹¹ http://www.informaticaumanistica.com/open-source/avindexer Accessed June 2021

¹² https://www.tiraccontolastoria.san.beniculturali.it/ Accessed June 2021

¹³ The foreseen address is: https://tiracconto.icbsa.it/



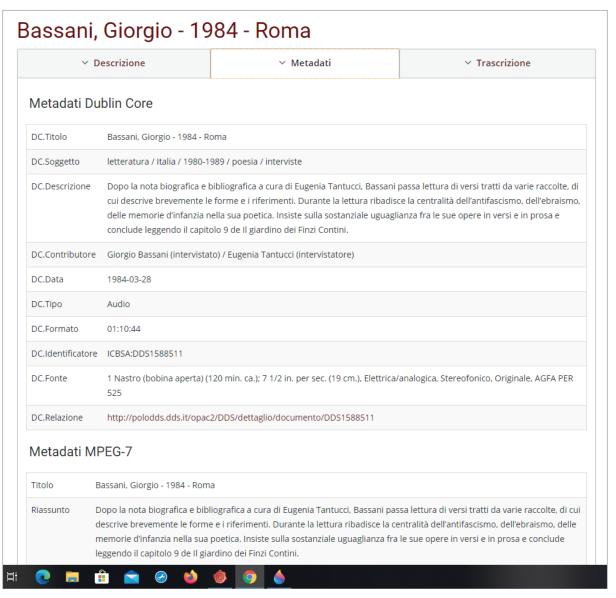


Fig. 5.



| Titolo | Bassani, Giorgio - 1984 - Roma |
|-------------------------|--|
| Riassunto | Dopo la nota biografica e bibliografica a cura di Eugenia Tantucci, Bassani passa lettura di versi tratti da varie raccolte, di cui descrive brevemente le forme e i riferimenti. Durante la lettura ribadisce la centralità dell'antifascismo, dell'ebraismo, delle memorie d'infanzia nella sua poetica. Insiste sulla sostanziale uguaglianza fra le sue opere in versi e in prosa e conclude leggendo il capitolo 9 de Il giardino dei Finzi Contini. |
| Produttore | Discoteca di Stato |
| Intervistatore | Tantucci Eugenia |
| Intervistato | Bassani Giorgio |
| Luogo | Roma |
| Data | 1984-03-28 |
| Genere | Intervista |
| Forma | Serie |
| Soggetto | letteratura / Italia / 1980-1989 / poesia / interviste |
| Lingua | it-IT |
| Durata | 01:10:44 |
| Parole chiave (SKOS) | Bologna / Università degli Studi di Bologna / Longhi, Roberto / Tantucci, Eugenia / Famiglie / Ferrara / Città / Ispirazione poetica / Attaccamento / Letteratura / Politica / Attività clandestina / Giovani / Intellettuali / Partito d'azione / Arresto / Antifascismo / Armistizio dell'8 settembre <1943> / Roma / Pubblicazione / Poesia / Bassani, Giorgio, Storie dei poveri amanti e altri versi / Bassani, Giorgio, Te lucis ante / Bassani, Giorgio, Un'altra libertà / Bassani, Giorgio, L'alba ai vetri / Redattori / Periodici / Botteghe oscure <periodico> / Paragone <periodico> / Bassani, Giorgio, Cinque storie ferraresi / Bassani, Giorgio, Gli occhiali d'oro / Bassani, Giorgio, Il giardino dei Finzi-Contini / Premio Viareggio / Bassani, Giorgio, Dietro la porta / Bassani, Giorgio, L'airone / Premio Campiello / Bassani, Giorgio, Le parole preparate e altri scritti di letteratura / Bassani, Giorgio, L'odore del fieno / Racconti / Saggi / Bassani, Giorgio, Epitaffio / Bassani, Giorgio, In gran segreto / Bassani, Giorgio, Il romanzo di Ferrara / Dediche / Bassani, Giorgio, In rima e senza / Premio Bagutta / Dortmund / Premio Nelly Sachs / Traduzioni / Narrativa / Sintassi / Lessico / Risentimento / Dolore / Vita / Storia / Solitudine / Discriminazione razziale / Religiosità / Laicismo / Emarginazione / Società / Contestazione / Dubbio / Antenati / Libertà / Bassani, Giorgio / Rima / Poeti / Versi / Alessandrini / Dodecasillabi / Settenari / Strofe / Iscrizioni / Epitaffi / Lapidi / Cimiteri / Case / Jahier. Piero /</periodico></periodico> |

Fig. 6.

JLIS.it vol. 13, no. 1 (January 2022) ISSN: 2038-1026 online Open access article licensed under CC-BY DOI: 10.4403/jlis.it-12731



References

Artini, Benelli and Melloni 2017. *Archivi di paglia. Gli archivi del distretto industriale della paglia in Toscana*, edited by Alessia Artini, Angelita Benelli, Silvia Melloni, Firenze: Edizioni Polistampa.

Capannelli, Emilio 2016. "Prime esperienze di collaborazione tra archivisti e bibliotecari" In *Il nome delle cose. Il linguaggio controllato come punto di incontro tra archivi, biblioteche e musei. L'esperienza del Gruppo linguaggi di MAB Toscana*, edited by Francesca Capetta, 17-20. Accessed June 2021. http://www.ilmondodegliarchivi.org/images/Quaderni/MdA Quaderni n1.pdf

Diadori, Pierangela 2021. "Ti racconto in italiano: voci del '900 per imparare l'italiano L2. Progetto ICBSA-UNISTRASI di didattizzazione di interviste sonore a personalità del '900." In La Nuova DITALS risponde 3, edited by Pierangela Diadori, Caterina Gennai, Elena Monami, in press. Roma: Edilingua.