OCLC Systems & Services: International digital library perspectives

Digitizing Oral History: Can You Hear the Difference?
Anthony Cocciolo

Article information:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

Digitizing Oral History: Can You Hear the Difference?

Introduction

For the last several years, my students from {remove name of institution for review purposes} MSLIS program have engaged in the digitization of oral histories ({references removed for review purposes}). Typically, the digitization activity is part of a larger effort to make oral histories available over the web, and usually involve students deploying a content management system, designing the front-end website, assigning metadata, and working within a rights framework. The oral histories come from a variety of archival institutions in the New York City area, including the Lesbian Herstory Archives, the Archives of the American Jewish Joint Distribution, Archives of the Center for Puerto Rican Studies at Hunter College, and the Archives of the American Field Service. These oral histories are most often contained on magnetic audiocassette, a once ubiquitous format now increasingly obscure.

As semesters progressed, I upgraded the digitization lab equipment to better adhere to professional audio digitization and archiving practices. Particularly salient practices are captured in *IASA-TC 04: Guidelines on the Production and Preservation of Digital Audio Objects* (2009), as well as the work of Casey and Gordon (2007). For example, each audio digitization workstation within the classroom (there are four in total) was upgraded to include a high-quality analog-to-digital converter (the ADC Benchmark USB[1]). These digital converters allowed for the creation of audio files at the rate and bit-depth recommended by audio archivists: 24-bits stored 96,000 times per second (or 96 kHz). Other upgrades include the setup of dual sets of

---

[1] http://www.benchmarkmedia.com/adc/adc1-usb

headphones per workstation so that students could verify that the audio content being played back corresponded exactly with what was being digitized.

Through these class projects, the switch from CD quality audio (16-bit/44.1kHz) to the accepted archival audio bit-depth and sample frequency (24-bit/96kHz) necessarily resulted in files approximately 3.3 times as large. For example, the digitization of one side of a tape in CD quality format usually results in files around 465 MB, and at archival quality around 1.5 GB. This increase in files size was manageable, but I did worry about the ability of the archival institutions I was partnering with—especially with ones with nothing other than grassroots support—to maintain over the long-term digital copies of audio tapes that were 3 GB each. Was the quality of the digital reproductions worth the tripled file size?

To research this question, I turned the issue over to my class by posing the following research questions:

RQ1 - Can MSLIS students discern the difference between oral histories digitized at archival quality (96 kHz/24-bit) versus CD-quality (44.1kHz/16-bit)?

RQ2 - Additionally, how important do they believe this difference is?

I choose my students as the research subjects not only for convenience sake, but more importantly for two reasons. First, a majority of the average student is under age thirty, which means they are not as subject to loss of hearing as older adults. There is well-documented

evidence that adults lose their ability to hear higher frequency sounds as they age, which will be discussed more thoroughly in the literature review. The second reason they were chosen is that as emerging librarians, archivists, and information professionals who have voluntarily chosen to take a course on digital archives, they are more committed to the preservation of historic material than the average adult. Hence, they are more likely to expend effort in deciding what is best for both the collection and the archival institution being served.

Before the research questions are directly addressed, relevant literature related to the digitization of oral history, as well as psychoacoustics, will be introduced. This will be followed by the study methodology, results, and conclusions.

Literature Review

Oral History and Digitization

Oral history "collects memories and personal commentaries of historical significance through recorded interviews," which then get "transcribed, summarized or indexed and then placed in a library or archives" (Ritche, 2003, p. 19). Frisch (1990) observes that oral history is "a powerful tool for discovering, exploring and evaluating the nature of the process of historical memory— how people make sense of their past, how they connect individual experience and its social contexts and how the past becomes part of the present, and how people use it to interpret their lives and the world around them" (p. 188).

For libraries and archives, the faithful, accurate and authentic reproduction of oral histories and sound recordings is particularly important because these aspects could influence the meanings perceived by future researchers. For example, if a researcher has reason to believe that a recording is not faithful to the original source (e.g., Why did the sound cut-out? What might be missing?) may cause a researcher to loose trust in the source and decide not to use it in her research. Similarly, it could influence researcher interpretations of the persons encoded on the recording (e.g., Why does his voice sound so different from other recordings? Is that really him speaking? Was he really that gregarious? Or dull? Why did people think she was such a great singer?). Thus, much of the value of a primary source such as a sound recording derives from its integrity, which also influences the researcher's interpretations of the subjects found a recording.

Through the 1990s, recording oral histories on magnetic audiocassette tape was fairly standard practice as evidenced by the array of archival institutions that hold oral histories in this format (e.g., Weig et al., 2007). Today, analog recording technology is considered obsolete and has been replaced by digital technology for both production and preservation (Alten, 2011; Casey and Gordon, 2007). It should be acknowledged that there are individuals who record new music on analog equipment for aesthetic purposes (Rudser, 2011), although there are no known examples of individuals continuing to use this technology for recording oral histories. With respect to new oral histories, recordings are most often captured using digital technology, such as audio or video recorders with solid-state memory cards. For older oral histories contained on an analog medium, the recommended best practice is to transfer the recordings to a digital format and preserve the original carrier (Casey and Gordon, 2007). Original carriers are best stored in

cool and dry environments, and the digital files are best stored in trustworthy digital repositories (Casey & Cordon, 2007; RLG & National Archives, 2008).

Audio archivists have developed well-accepted practices for digitization of sound recordings, which is best captured in *IASA-TC 04: Guidelines on the Production and Preservation of Digital Audio Objects* (IASA, 2009), as well as the work of Casey and Gordon (2007). However, relatively little research has been conducted that illustrates how average users perceive the difference between reproductions created using archival audio standards versus lesser standards. Related research has explored minimum standards for digitizing speech-based recordings on audiocassette (e.g., 48kHz/24-bit), which could be useful for smaller organizations where the cost of digitization and file maintenance is too great (Jackson, 2013). To address the issue of what individuals hear, relevant research from psychoacoustics will be introduced.

Psychoacoustics

Psychoacoustics research aims to "determine the relation between the physical stimuli (sounds) and the sensations produced in the listener" (Plack, 2005, p. 4). With respect to psychoacoustics and sound reproduction systems, one may assume that humans would prefer the highest fidelity—or most accurate—reproduction. However, past research has demonstrated that listeners do not necessarily prefer listening to the highest fidelity sound recordings. As early as 1956 Kirk demonstrated that learning and past listening experiences help define listening preferences. In studying 210 college students, he found that "average college student prefers music and speech reproduced over a restricted frequency range rather than an unrestricted

frequency range" (p. 1113). This research indicates that particular sonic profiles may get associated with positive emotions produced not only by the sounds themselves but by the related listening experiences (e.g., pleasurable moments with friends), causing listeners to prefer these sonic profiles over sonic profiles that offer higher fidelity. This phenomenon has been attributed to the preference among some contemporary college age students for sound recordings subjected to lossy compression, where a perceptual sound encoder (such as the MP3 encoder) removes sounds that humans should not be able to hear (e.g., one sound masks another) (Sterne, 2012). Newspapers and magazines have picked-up related research from music scholars such as Jonathan Berger, which have resulted in headlines such as "Young music fans deaf to iPod's limitations," and "Are iPods killing music perception?" (Ahmed and Burgess, 2009; LeFevre, 2009). However, this point continues to be debated as new research suggests that teenagers may prefer the higher fidelity recordings (Olive, 2011).

Linking favorable experiences to a particular sonic profile could explain the preference some have for analog recordings on magnetic tapes or vinyl records. Despite the observation that analog recordings produce a higher signal-to-noise ratio than digital recordings (e.g., tape hiss or crackle of vinyl record), some listeners continue to prefer these sonic profiles (e.g., Felten, 2012).

In addition to learning experiences that determine listening preferences, other factors such as physiological factors contribute to what individuals hear. As mentioned earlier, one particularly salient aspect is age, where hearing loss is "a very common problem affecting older adults" (Cruickshanks et al., 1998, p. 879). Alten (2011) notes that as "gradual deterioration of the

auditory nerve endings occurs with aging, it usually results in a gradual loss of hearing first in the mid-high-frequency range, at around 3,000 to 6,000 Hz, then in the lower-pitched sound" (p. 17). Thus, after age 55 there is an accelerating rate of loss (Patterson et al., 1982). Listener age has also been shown to affect the ability to hear speech in noisy environments (Dubno et al., 1984).

In sum, the research from psychoacoustics indicates that there is not a single or ideal way to reproduce sound, and reproduction preferences are in-part determined by past experiences and learning, the content itself and physiology.

Methodology

Study participants listened to three sets of oral histories clipped to two minutes in duration. The clips were created from oral histories that were being digitized in class during the Fall 2013 and Spring 2014 semesters across five courses being taught by the researcher (three sections of LIS 665 Projects in Digital Archives, and two sections of LIS 668 Projects in Moving Image and Sound Archiving). The researcher created these clips by playing back the first two minutes of an audiocassette tape and digitizing at 96 kHz/24-bit. The tape player used is an Alesis Tape2USB tape player connected to an ADC Benchmark USB analog-to-digital converter. The file is saved as a 24-bit WAV file using the open-source software program Audacity.[2] A copy of the WAV file was down sampled to 44.1 kHz / 16-bit using the Windows program r8brian, with the

---

[2] http://audacity.sourceforge.net/

conversion option for "Highest Quality" down-sample conversion.[3] The oral histories used to create the two-minute clips include:

1) An interview by dance critic Barbara Newman with dance choreographer Mark Morris from August 15, 1996. This tape is part of a personal collection of interviews with dancers and choreographers that formed the basis of her book, *Grace Under Pressure: Passing Dance Through Time.*

    File Size Information: Archival: 70 MB, CD-quality: 21.4 MB

2) An interview by Elizabeth Kennedy with oral history subjects Bobbi and Terri (fictitious names) for the Buffalo Women's Oral History project. These oral histories formed the ethnographic dataset for Kennedy and Davis' study of lesbian women in Buffalo from the 1930s-1950s, resulting in the seminal LGBT studies text *Boots of Leather, Slippers of Gold: The History of a Lesbian Community.* This interview was from September 25, 1982. This tape is held at the Lesbian Herstory Archives in Brooklyn, NY.

    File Size Information: Archival: 69.5 MB, CD-quality: 21.3 MB

3) An interview with Del Martin and Phyllis Lyon from May 9, 1987. Martin and Lyon formed the first lesbian political and social organization in the United States in 1955. The group was formed in San Francisco and named the Daughters of Bilitis. This tape is also held at the Lesbian Herstory Archives in Brooklyn, NY.

    File Size Information: Archival: 69.2 MB, CD-quality: 21.2 MB

---

[3] http://www.voxengo.com/product/r8brain/

The two versions of the recordings were randomly assigned the letter "A" or "B," and students were asked to play back both "A" and "B" and decide which was the "archival quality" (or 24-bit / 96 kHz), and which was the "CD quality" (or 16-bit / 44.1 kHz). All six files were placed on the desktop of a Windows computer and grouped together visually. Each workstation in the computer lab connected with a pair of inexpensive JVC ear bud headphones that had been cleaned by the researcher using an anti-bacterial cloth. The same brand and model of headphones were used by all participants to eliminate any differences introduced by varieties of headphones. The student was asked not to inspect any of the metadata related to the file, such as file size, which would give away which recording was which. The students were also let know that that they could re-listen to the recordings, rewind, and compare as much as needed to reach their determination. The students would fill-out the survey included in the appendix and return it to the researcher. The research was conducted across 5 class sessions in October, 2013 and February 2014.

The six files used for this study are available for listening and download.[4]

The participants in the study represent graduate students in a MSLIS program. 53 individuals participated in this study, with an average age of 30.2 (standard deviation of 7.9). The oldest participant was 58 years old and the youngest was 23 years old. 79% of participants are female, and 21% male.

Results

---

[4] {URL removed for review purposes}

RQ1 - Can MSLIS students discern the difference between oral histories digitized at archival quality (96 kHz/24-bit) versus CD-quality (44.1kHz/16-bit)?

Based on the responses of 53 MSLIS students, Table 1 reveals that less than half of the time on average could students discern which was the archival quality versus the CD quality recording. However, this was dependent on the actual recording, which is evident from observing that Test A had over sixty perfect correct identification, and B and C were only correctly identified around 35 percent of the time.

Table 1. Percentage of sound recordings correctly identified as being "archival" digitization versus "CD-quality" digitization

{Insert Table 1 Here}

RQ2 - Additionally, how important do they believe this difference is?

Participants were asked, given the sound test they just listened to and if you were digitizing an important oral history collection, how important is the difference between CD quality (16-bit / 44.1 kHz) and archival quality (24-bit / 96 kHZ) digitized sound? They answered this question on the scale (0 = Not at all important, 1 = A little bit important, 2= Important, 3 = Very important). The mean response was 1.3 (standard deviation of 0.78), with the most frequent response being "A little bit important" (31 individuals marked this response).

The researcher opened up a discussion after the surveys were submitted about how important the difference in quality was. One student mentioned that there was a slight difference in what you can hear in terms of the background noise between the archival and CD-quality version. The archival version made clearer the commotion in the background; however, she noted there was no difference in what could be heard from the primary speakers in the oral histories. She said that if the purpose of an oral history is to record the vocalized memories of the speaker, and the background noise is not the primary concern (as opposed to another kind of recording that might try to capture the noise of a landscape), then the quality difference was of little importance. One student mentioned that the sound of the tape hiss was slightly different between the two recordings; however, this was of little importance to her.

Discussion and Limitations

Results from this study reveal that MSLIS students have difficulty discerning the difference between oral histories digitized at archival quality (24-bit / 96 kHz) and CD quality (16-bit / 44.1 kHz). However, this can vary to some extent based on the actual content of the recording. Once completing this discernment test, the students most often found the difference between the two formats as being a "little bit important." However, there were a minority of students—16 in total—who though the difference was important or very important.

One limitation of this study was that the students all used inexpensive ear bud headphones in a classroom within a less than ideal sound digitization environment. Higher-quality headphones that encompassed the entire ear could possibly reveal more details. And although the classroom

was relatively quiet during the playback of the tests, some background noise is always present in the classroom's Manhattan-based environment (e.g., street noise, noise from floors above and below). With respect to playback environment, Casey and Gordon (2007) recommend that "preservation transfer work is best undertaken in a studio designed as a critical listening space" (p. 10). At a minimum, the playback "studio must at least be free from ambient noise, it must be removed from other work areas and traffic, and its acoustic weaknesses should be well understood." (p. 10). While although the facilities for the sound test and inexpensive headphones were chosen because they were the only resources available the researcher, they are appropriate choices because they are not dissimilar to the environment where most researchers will listen to sound recordings. For example, inexpensive headphones are in wide use, such as the white ear buds that come pre-package with the Apple iPhone, and library reading rooms and research spaces are filled with ambient noise.

A final limitation is that the tape deck used to playback the original media is a consumer-grade tape deck of recent vintage. Most audio archivists recommend using cleaned and restored professional grade equipment (Casey and Gordon, 2007; Jackson, 2013). For example, consumer grade equipment does not allow adjustment to the azimuth, which is the angle at which the record/playback head connects with the tape. Unfortunately, this equipment can only be bought used, and it is difficult to purchase such equipment within the confines of contemporary higher education purchasing practices, which prefer to purchase new equipment from select retailers, and shy away from allowing purchases from less well-worn paths.

Conclusion and Implications

A major implication of this study is that most untrained listeners may not be able to tell the difference between archival quality and CD quality digitization. If MSLIS students—who as part of their coursework have received training in the best practice in audio digitization—conclude on average that it is "a little bit important," it is likely that the general public may feel similarly or even more extremely. For archival institutions, this could complicate justifying the additional digitization expense as well as increased file size. However, this is not to suggest that archivists abandon well-established sound digitization practices that produce results that audio archivists (and those able to hear fine grain audio differences) find superior. Rather, it does imply that additional work may be needed to train listeners to discern these fine grain differences, and appreciate the highest-fidelity replication of original audio recordings. This is no easy task; however, some listening education could help. For example, a series of exercises could be designed where important details that can only be revealed through higher fidelity recordings—and masked through lesser quality recordings—could help make the point of maintaining as much of the original audio content through digitization as possible.

References

Alten, S. R. (2011), *Audio in Media,* 9th Edition, Wadsworth, Boston, MA.

Ahmed, M. and Burgess, K. (2009), "Young Music Fans Deaf to iPod's Limitations." *London Times Online*, 5 March 5, available at: http://www.thetimes.co.uk/tto/technology/gadgets/article1860325.ece (accessed 7 December 2013).

Casey, M. and Gordon, B. (2007), *Sound Directions: Best Practices for Audio Preservation*, Harvard University and Indiana University, Cambridge, MA and Bloomington, IN.

Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E. K., Klein, R., Mares-Perlman, J. A., Nondahl, D. M. (1998), "Prevalence of Hearing Loss in Older Adults in Beaver Dam, Wisconsin: The Epidemiology of Hearing Loss Study", *American Journal of Epidemiology,* Vol. 148 No. 9, pp. 879-886.

Dubno, J. R., Dirks, D. D., Morgan, D. E. (1984), "Effects of age and mild hearing loss on speech recognition in noise", *Journal of the Acoustical Society of America,* Vol. 76 No. 1., pp. 87-96.

Felton, E. (2012), "It's Alive! Vinyl Makes a Comeback", *Wall Street Journal*, 27 January, available at: http://on.wsj.com/x5SrGy (accessed 22 January 2014).

Frisch, M. (1990), *A Shared Authority: Essays on the Craft and Meaning of Oral and Public History.* State University of New York Press, Albany, NY.

IASA Technical Committee (2009), *Guidelines on the Production and Preservation of Digital Audio Objects*, 2nd Edition, edited by Kevin Bradley, available at: http://www.iasa-web.org/tc04/audio-preservation (accessed 22 January 2014).

Jackson, D. J. (2013), "Defining Minumum Standards for the Digitization of Speech Recrodings on Audio Compact Cassettes", *Digital Preservation, Technology & Culture,* Vol. 42 Num. 2, pp. 87-98.

LeFevre, T. (2009), "Are iPods killing music perception?", *GizMag*, 18 March, available at: http://www.gizmag.com/ipods-killing-music/11236/ (accessed 7 December 2013).

Kirk, R. E. (1956), "Learning, a Major Factor Influencing Preferences for High-Fidelity Reproducing System", *Journal of the Acoustical Society of America,* Vol. 28 No. 6, pp. 1113-1116.

Olive, S. (2011), "Some New Evidence That Teenagers May Prefer Accurate Sound Reproduction", conference paper presented at Audio Engineering Society Convention 20-23 October, New York, NY.

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., Milroy, R. (1982), "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold", *Journal of the Acoustical Society of America*, Vol. 72 No. 6, pp. 1788-1803.

Plack, C. J. (2005), *The Sense of Hearing,* Psychology Press, New York, NY.

Research Libraries Group and National Archives and Records Administration (2008), *Trustworthy Repositories audit and certification: criteria and checklist*, available at: http://wiki.digitalrepositoryauditandcertification.org/pub/Main/ReferenceInputDocuments/trac.pdf (accessed 22 January 2014).

Ritchie, D. A. (2003), *Doing Oral History: A Practical Guide*, 2nd Edition. Oxford University Press, New York, NY.

Rudser, L. (2011), "Miss the Hiss? Fanatics Flip for Tunes on Cassette Tapes", *Wall Street Journal,* 21 October, available at: http://on.wsj.com/nMyBZB (accessed 15 January 2013).

Sterne, J. (2012), *MP3: The Meaning of a Format*, Duke University Press, Durham, NC.

Weig, E, Kopanna, T. and Lybarger, K. (2007), "Large Scale Digitization of Oral History: A Case Study", *D-Lib Magazine,* Vol. 13. Num. 5/6, available at: http://www.dlib.org/dlib/may07/weig/05weig.html (accessed 7 December 2013).

Biographical Details

Anthony Cocciolo is an Assistant Professor at Pratt Institute School of Information and Library Science, where his research and teaching are in the areas of digital archives, moving image and sound archives, and digital libraries. He completed his doctorate from the Communication, Computing, Technology in Education program at Teachers College, Columbia University. Prior to Pratt, he was the Head of Technology for the Gottesman Libraries at Teachers College, Columbia University.

Appendix

Survey: Digitizing Oral Histories: Can you hear the difference?

Test 1: Which sound file is the archival quality (24-bit/96 kHz) sound file (as opposed to the CD quality (16-bit/44.1 kHz))? (Circle one answer)

A   B   I can't tell the difference

Test 2: Which sound file is the archival quality (24-bit/96 kHz) sound file (as opposed to the CD quality (16-bit/44.1 kHz))? (Circle one answer)

A   B   I can't tell the difference

Test 3: Which sound file is the archival quality (24-bit/96 kHz) sound file (as opposed to the CD quality (16-bit/44.1 kHz))? (Circle one answer)

A   B   I can't tell the difference

Given this sound test, and if you were digitizing an important oral history collection, how important is the difference between CD quality (16-bit / 44.1 kHz) and archival quality (24-bit / 96 kHZ) digitized sound? (Circle one answer)

Not at all important
A little bit important
Important
Very important

What year were you born in?
19_____

What is your gender? (Circle answer)
Male
Female

Table 1. Percentage of sound recordings correctly identified as being "archival"

digitization versus "CD-quality" digitization

Test	% of participants correctly identified the archival quality digitization
A	62.3%
B	34.0%
C	37.7%
Overall	**44.7%**