

Creating Workflows for Oral History Transcription and Curation: A Case Study Involving Oral Histories in Satellite Meteorology

Aaron Gregg

Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI

Jinny Nathans

American Meteorological Society, Boston, MA

Jean M. Phillips

Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI

Sophie Mankins

American Meteorological Society, Boston, MA

Katherine Johnson

Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI

Abstract

This article presents a case study of the transcription and curation of a set of oral histories conducted by the American Meteorological Society (AMS) in collaboration with the Schwerdtfeger Library at the University of Wisconsin-Madison Space Science and Engineering Center. The oral histories will become part of the AMS Oral History Project and the Schwerdtfeger Library's Digital Collections. Oral history interviews were conducted with prominent figures in satellite meteorology at the AMS Joint Satellite Conference in Boston, Massachusetts in 2019. These interviews invited leading remote sensing scientists to share perspectives on their careers, the innovations in and evolution of satellite capabilities and intersections with their careers, as well as future directions of the field. They elaborated at length on subjective details of their lives and work that might otherwise not have been captured in scholarly literature or in other locations. Interview transcripts were processed to conform to the Smithsonian's Oral History Program Style Guide and were archived according to the Oral History Association's Manual of Best Practices for Archiving. This paper seeks to examine the processes involved in transcribing and preparing files for curation as a case study for other institutions considering similar projects in the sciences.

Article Type: Case study

Introduction

This article describes an oral history project undertaken as a joint effort between the

American Meteorological Society and the University of Wisconsin-Madison Schwerdtfeger Library. Senior librarians brought experience conducting oral history interviews to the project,



Endnotes: The Journal of the New Members Round Table
Volume 10, Issue 1, 2021
pp. 13-20
Published by the American Library Association
© Retained by the Authors
ISSN: 2159-0591

having undertaken similar projects in the past for their respective organizations. The lead author, new to the field, did not have prior experience in oral history work, but was tasked with developing a workflow for the subsequent transcription and audit-editing of interviews within this environment. As such, this article offers a perspective on conducting oral history work by librarians with a range of prior experience doing so—from limited prior experience to greater experience—with guidance from a body of literature and established methods.

This article describes the methodologies for oral history production used and developed within this particular project and was written to serve as a potential model for other organizations interested in undertaking oral history projects, as a way to provide further context, guidance, and practical considerations. It examines the technical aspects of oral history production, with a focus on transcription processing. Specifically, it breaks down observed processing needs for transcripts and how they were addressed systematically within this project. Furthermore, it presents how conventions can be created and structured while drawing from recommended best practices according to a chosen style guide.

The oral history project described in this article was conducted for the scientific field of satellite meteorology. As such, this article also represents a perspective on conducting oral history work within science by providing direction on transcribing field-specific jargon or acronyms. The article touches on the subjective nature of oral history and, consequently, the additional viewpoint it contributes to research.

Literature Review

Oral History as a Methodology

There is considerable literature on oral history as a methodology. Some notable guides include Donald A. Ritchie's *Doing Oral History* (2003) and Nancy McKay's *Curating Oral Histories: From Interview to Archive* (2016). McKay presents oral history interviewing as a historical methodology in which an interviewer poses questions for a narrator, who then provides and

controls the content of the interview and in doing so serves as its primary creator. This is a collaborative process that involves dialogue throughout, which is one aspect that makes oral history interviews different from many types of static historical documents produced by individuals (Abrams, 2016). In addition, oral history interviews typically seek to elicit more subjective and unfiltered accounts in contrast to many types of officially published works that are deliberately presented in more depersonalized and abbreviated forms (Weiner, 1998). Rather than merely providing depersonalized facts, oral history interviews allow for the exploration of an individual's own experiences as well as insight into their thoughts and motivations (Abrams, 2016).

Oral history interviews have been conducted in a wide range of contexts. Many oral histories have been conducted to capture historical records in sociological or anthropological contexts, where, as noted by Abrams (2016), they often serve as a way to present voices that may not have been heard or preserved otherwise. Oral histories have also been conducted to capture historical accounts of science, but these, in contrast, often come from individuals who have already left large paper trails of articles or books, given speeches, or left a large volume of personal records or correspondences, as noted by Weiner (1998). Despite this, oral histories of science can bring out other types of historical information which may not be typically captured or well-represented in those other places or formats. For instance, oral history as a mode has the ability to reveal more of the institutional or ideological factors behind the contributions of particular scientists, or more about how details of their personal lives drove or affected their work (Weiner, 1988).

This project aimed to capture oral histories for the scientific field of satellite meteorology and the literature was searched for accounts of similar oral history projects for the history of science. The search uncovered similar oral history work with major figures in the field of planetary science as well as larger scale oral history collection within the American Institute of

Physics (Sears, 2012) and the Lamont-Doherty Earth Observatory (Levin & Doel, 2000).

Oral History Creation and Processing

Oral histories are unique relative to other types of content in terms of how they are produced and curated. As described by Powers (2005) and others, modern oral history creation typically involves the recording of the audio content of an interview as well as the subsequent production of a transcript based on that audio. In modern oral history, the interview audio recording is considered the primary document (McKay, 2016), while the derived transcript is considered a distinct type of work. In this sense, the production of a derived transcript is, to an extent, an act of creation on its own. Furthermore, the process of creating the transcript requires some considerations and stages of work that are outside of typical workflows for other types of content.

The preparation of a transcript typically involves several successive processes. First, the verbal content of an interview must be transcribed in order to produce a verbatim transcript. A first draft can be achieved with voice recognition software; however, past attempts with available software, such as those described by Yaco (2007) and Ciota (2019), have been found to result in relatively incomplete and inaccurate transcripts, necessitating completion by a human listener. Audit-editing then involves replaying a recording while monitoring the transcript in order to assess consistent correspondence and correct inaccuracies (Ritchie, 2003). Processing of audio into verbatim transcripts involves further decisions about what to include as words versus non-words, decisions about punctuation, and decisions about how the transcript is represented for overall consistency, and in that sense is a non-agnostic form of editing. Furthermore, there is routinely a trade-off available between preserving consistency with the original audio versus making further edits which would improve readability or grammatical correctness but at the cost of that consistency (Powers, 2005). Transcription can also involve the insertion of additional notation to convey meaning, such as indications of laughter—laughter in particular is a

paralinguistic feature that is often well-represented in interview transcripts (Myers & Lampropolou, 2015).

Another facet of oral history transcription is the role of memory and attention to detail in the process. Researchers in the field of psychology have done considerable exploration of the concepts of attention as well as working memory. The latter has been conceptualized as a system for the temporary storage and manipulation of information held in mind that has a limited capacity (Baddeley, 2010). Transcribing and audit-editing transcripts are activities which place significant loads on attention and demands on working memory; when listening to an audio segment in order to produce and edit its transcript, there are a number of pieces and types of auditory information coming across in succession or simultaneously as well as a number of rules that the transcriber is aiming to apply to process those pieces. Because attentional and working memory capacities are so limited, it is challenging to maintain awareness and monitoring in order to “catch” all of the relevant types of content so that they can be accurately represented in the final transcript.

Oral History Work in Libraries and Archives

In the literature, there are a handful of published case studies of libraries and archives conducting oral history programs. These accounts touch on various aspects of the entire oral history production and curation process, including some initial interview considerations, basic aspects of transcription, as well as some post-production considerations such as metadata and content hosting.

In terms of transcription, previous case studies presented several relevant themes. First, transcribing and editing is time- and labor-intensive, and this time is extended by processing steps such as researching proper names (McKay, 2016). This aspect was commonly mentioned in case studies (e.g., Boutin-Cooper, 2019). There are outsourcing options in the form of professional transcription or editing services, but they are often cost-prohibitive (Hurford & Read, 2008). For this

reason, organizations creating transcripts in-house commonly hire student hourly workers to transcribe recordings (e.g., Wright, 2011). As for specific issues in transcription, accidental misinterpretation of homophones can be an issue, but beyond that, transcribers are sometimes unfamiliar with certain references made by interviewees (e.g., Ciota, 2019). Additionally, Humphries (2015) mentions how speech naturally contains many fragments and trailing sentences, which necessitates decisions surrounding their treatment. As for editorial consistency and standardization, clear reference to the use of external style guides or internally developed standards was somewhat limited in the literature. The Middletown Digital Oral History Project's case study (Hurford & Read, 2008) mentioned the use of an external style guide from the Baylor University Institute for Oral History, as one example. However, some other relevant case studies did not elaborate on editorial standards, and it was not always clear which, or whether, an external style guide was adopted or followed. Generally speaking, existing case studies from libraries and archives touched on, but didn't go into significantly greater depth on, the transcription portion of oral history production. This paper will emphasize the transcription portion of oral history production in order to further contribute to the available literature.

Aims of this Project

Background

The Schwerdtfeger Library is a special library that is embedded within the Space Science and Engineering Center (SSEC) at the University of Wisconsin-Madison. Its mission is to support research, teaching, and learning needs of SSEC, affiliated researchers, the University's atmospheric and oceanic sciences department, as well as the broader campus and community. Important within this context, the library curates and hosts special collections in multiple formats, including print and digital, related to the fields of atmospheric science and satellite meteorology. The American Meteorological Society (AMS) is a professional organization devoted to the sharing of information within the atmospheric and related

sciences (American Meteorological Society, 2021), including archiving and preserving materials relevant to these areas of science and associated technologies.

In 2019, the AMS hosted a joint conference with NOAA (the National Ocean and Atmospheric Administration) and EUMETSAT (the European Organisation for the Exploitation of Meteorological Satellites; American Meteorological Society, 2019) that brought together international scientists and researchers in the field of satellite meteorology.

This particular conference was held two months ahead of the 100th anniversary of the AMS, providing a rich opportunity to reflect back on the history and growth of the field of satellite meteorology. At just over 50 years old, the field is relatively young, to the extent that a number of those in attendance at this conference were pioneers in their own right, having led or collaborated on many of the advances in the field over the past four to five decades, including through leading key organizations. However, the conference also came in an era of major improvements in satellite technology and computing infrastructure, so it provided an opportunity to look forward to upcoming developments and the future of the field.

Librarians from the AMS and the Schwerdtfeger Library took the opportunity of this collective conference to conduct oral history interviews with notable attendees who had spent their careers in this field. These interviews explored past developments as well as current trends and emerging opportunities in satellite meteorology. They also provided an opportunity for interviewees to elaborate on their lives, perspectives, and experiences.

Interviews were recorded, and through the joint effort of the AMS and Schwerdtfeger Library, they were subsequently transcribed. They were then curated for archival use within both the American Meteorological Society's Archive and the Schwerdtfeger Library's Digital Collections. The audio contents of these interviews and their associated transcripts will be preserved and made publicly available by these two

organizations in order to provide broad access to these important historical resources.

Implementation and Workflows

Preparation and Interviewing. In consultation with senior leadership at SSEC, interviewers arrived at a list of 22 leaders in the field of satellite meteorology who could be invited to sit for an interview, many of whom had been working in the field for up to 45 years. These initial conversations also included perspectives on roles and contributions of each candidate. Of the 22, 14 were planning to attend the conference and were available for a 30- to 40-minute interview. Additional background preparation included: 1) requesting a curriculum vitae from each candidate to gain a sense of the scope of their careers, 2) reading relevant peer-reviewed articles, 3) reading other articles about the candidates published by their organizations or media outlets, and 4) developing questions.

Prior to the interview, candidates were provided with a set of general expectations for what the interview would entail. Interviewees were invited to specifically speak on a) their decision to become a meteorologist and why they decided to pursue that as a career, b) a professor or mentor who made a difference in their undergraduate or graduate education, c) the evolution of satellite capabilities and its intersection with their career, and d) their perspective as a remote sensing or satellite scientist and where future systems are headed. They were invited to share stories about their education, career, research, or any other facet of their experiences as a meteorologist and were informed that their perspectives on their career and innovations in the prior decades would come to serve as a vital primary resource for the history of meteorology.

During the interviews, participants were very open and eager to share stories about their early experiences, working relationships, and career directions. The process allowed for individual and personal stories and perspectives to emerge. For instance, as a route into meteorology, several interviewees had flown as pilots, which also gave them perspectives on how different groups communicate and produce

tools that are useful for each other. Many other interviewees talked about public engagement with the science of meteorology, as well as the continued benefits of widespread data sharing. Interviewees touched on changing fields, working across disciplines, and dynamics of organizational and international cooperation.

The ability to share stories in this setting also gave way to some particularly interesting and humanizing personal anecdotes. For example, one participant (who years later went on to become the Director-General of Europe's EUMETSAT) mentioned flipping a coin when deciding which field to pursue. Another participant mentioned talking with a future collaborator on a long train ride early in his career. Others mentioned situations wherein they and others had to regroup and persevere after instrument and launch failures, and the trials, emotions, and ultimate pride associated with these events.

Transcription and Processing. From a technological standpoint, interviews were recorded with a bi-directional microphone using Audacity, an open-source and cross-platform audio software application. The files were saved in the preferred Waveform Audio File Format (.wav). Audio segments were initially converted to text using voice recognition software, in this case Dragon software, but this step was insufficient to produce complete and accurate transcripts. Initial outputs were missing numerous words, and many words were transcribed incorrectly. In addition, the software was unable to transcribe acronyms and certain instances of technical vocabulary. Additional processing was required to bring transcripts into conformity with selected standards.

In manual transcription of the interviews, a logical workflow had to be developed. Following review of some existing guides, we adopted the Smithsonian's Oral History Program Style Guide. It was helpful in providing an underlying framework for consistency within and between transcripts. By adopting a specific style, many processing conventions could be effectively planned at the outset. Nevertheless, other issues

arose and were solved based on needs and experience working with the transcripts.

Given the demands of transcription on attention, breaking down or separating different aspects of processing can be an effective strategy in transcribing and editing and was found to be an effective approach within this particular project. For example, a transcriber can: a) listen for a selective set of aspects during a pass and b) in separate batches, handle well-defined aspects of the transcript that do not need concurrent audio play-back to complete. Additionally, a transcriber can: c) process a separate subset of defined aspects during distinct listening if there is time available to do so, which was done to an extent in this project. These divisions were initially planned by taking the plain text of the selected style guide and sorting it into rough categories based on the type of processing required. Many requirements and rule-based actions could be quickly sorted into groups based on those that required audio-playback versus those that could be completed without audio. They were then grouped into sets of similar actions that could be completed in batches. This approach expedited subsequent processing of the transcripts.

During each processing stage, conventions from the chosen style guide were applied at the relevant point. For instance, formatting conventions such as font, margins, and justification could be established at the outset, and bracketed notation for interruptions or inaudible segments could be included as encountered during listening. In the case of this project, initial transcription focused on capturing the words of the interview fully and accurately along with their associated sentence structures, including false start fragments as set off with em dashes and some other issues, such as adopting unambiguous punctuation and consistent notations of laughter. After initial listening, a handful of short steps to further refine the transcript were developed and undertaken separately. Representations of special cases such as spoken numbers, fractions, percentages, units of measurement, times, dates, and years were standardized to conform to the chosen style guide. Locations and titles were likewise systematically represented.

Acronyms and field-specific jargon were then searched and verified based on relevant websites or publications. Again, in accordance with the chosen style guide, acronyms were expanded with bracketed comments, a practice helpful for later listeners. In a similar fashion, names mentioned in the interviews were also searched independently in a batch to verify their spelling, often using websites or co-authorship on publications for verification. As a final step—and before sharing with each interviewee—other members of the group read the transcripts for spelling, punctuation, or other errors. These later processing actions could be accomplished, to some extent, independent of audio playback—though listening was necessary in certain cases. Furthermore, separating processing steps in this systematic way allowed the transcriber to focus on a smaller, more manageable number of concerns during initial listening and at each subsequent stage.

As for the most common issues encountered: Transcripts were intended to be a verbatim representation of the audio file as opposed to being further edited for grammatical correctness in ways that might diverge from the original audio. An ever-present challenge was proper sentence construction while respecting and working within the constraints of the original speech, which often was not as polished as edited, written language and can vary with the speaking style of individual participants. Overall, there was very low prevalence of problematic non-words such as “um,” but false starts of sentences were very common. Fillers and repetitions were commonplace but were largely retained in the transcript in order to remain faithful to the audio. Excessive background and low-volume affirmatives were predominantly left out in transcription, and this was an important choice because it allowed narratives to be presented with their full poignancy rather than being excessively disrupted on the page. Lastly, some small portions of the interviews remained inaudible because of participants coughing or speakers trailing off and remained marked as such.

Because of the technical nature of this particular set of interviews (on satellite meteorology and

some related topics such as physics or data science), technical vocabulary and acronyms were commonplace. Researching and verifying acronyms was a significant part of this project. Carefully checking them with context allowed for the correction of residual mistakes in immediate transcription—for example, the acronym CYGNSS (Cyclone Global Navigation Satellite System) was initially written as Cygnus, another specific name from elsewhere in the domain. Most technical references were, however, readily accessible to search, which made the process of finding and verifying them unproblematic. Additionally, many terms were mentioned across a number of interviews, so familiarity with them from transcribing initial interviews allowed for recognition as they re-occurred in subsequent interviews. As for personal names, some could not be immediately found online, but those that were likewise allowed correction of some names that would have been unwittingly transcribed incorrectly otherwise—for example, the last name “Bohren,” initially written as “Borin” or “Velden,” written as “Felden.”

After being prepared within the organization, but prior to publication, transcripts were shared with original interview participants. This was done out of respect and ethical consideration for the interviewees, and it also gave participants the chance to provide their input as to the spellings or meanings of any particular reference that may have been transcribed incorrectly.

Interviews will be included among the digital collections of the Schwerdtfeger Library and hosted separately by the AMS. In addition, the transcripts will be saved as searchable PDFs, allowing for full text searching of the documents.

Conclusions and Discussion

This oral history project builds on previous oral histories undertaken by the AMS and the Schwerdtfeger Library. Senior librarians brought prior experience with the oral history life cycle including recruiting participants and conducting interviews. Given this background along with the prominence of the AMS and UW-Madison in the field of satellite meteorology, the candidates trusted the process and provided rich accounts

of their lives and work. This mode of interviewing successfully gave attendees an opportunity to share subjective experiences and impressions of work in the field of satellite meteorology as well as tangentially related topics, pieces of which likely have not been captured in more official publications or in other places such as shorter-form interviews. They embraced the occasion, and their accounts gave rise to a number of interesting insights and anecdotes about their respective paths. Overall, the interviews resulted in rich personal perspectives on satellite meteorology and the lives of these individuals within it. They will join a larger body of research resources for future listeners, including historians, researchers, and others.

Within this context, the transcription portion was planned and undertaken by the lead author without prior experience in similar projects. Transcription presented some issues due to the nature of spoken word and the technical complexity of these particular interviews. However, these challenges largely occurred as some common types which could be readily recognized and treated categorically. Issues in interviews (such as false starts, filler words, etc.) tended to follow a small number of common patterns which could be approached similarly. These recurring issues were able to be addressed by developing a process to systematically segment, or categorize, aspects of the transcripts and process them accordingly. In transcription of oral histories from other fields, transcribers may encounter different processing needs depending on the technical complexity of the area. Given this, any processing stage related to verifying and elaborating on jargon may be a greater or lesser consideration than it was during this project. However, this is a crucial concern for oral histories of science in order to make them readily accessible to future audiences.

In this project, it was also helpful to outline conventions for transcription as much as possible at the outset, working from the Smithsonian’s Oral History Program Style Guide. A fundamental recommendation is to select an oral history program style guide and use its conventions consistently throughout the life cycle

of an oral history project. Organizations are encouraged to examine available style guides and choose the one that is best suited to their needs. Secondly, it was helpful to break up transcript processing into some well-defined discrete pieces. It is recommended that other libraries and library workers undertaking oral history programs for the first time approach them in a similar manner. Lastly, it is recommended that other libraries and archives create the opportunity to undertake oral history projects. Our organizations are rich with history, stories, perspectives, and personalities that are not necessarily represented in the peer-reviewed literature but are important to creating a fuller picture of a field at a given time, and oral history provides both an excellent window into that history and a representative way of preserving it for the future.

Acknowledgements

The authors would like to acknowledge the enduring contributions of the late AMS Archivist Jinny Nathans to this oral history project and to the field of atmospheric science librarianship. In addition, we thank the AMS and the Space Science and Engineering Center for support for the project.

References

- Abrams, L. (2016). *Oral history theory* (2nd ed.). Routledge.
- Baddeley, A. (2010). Working memory. *Current Biology*, 20(4), 136-140.
- Boutin-Cooper, A. (2019). Oral History as Care: Preserving memories and maintaining stakeholder relationships. *Visual Resources Association Bulletin*, 46(2), 1–8.
- Ciota, R. (2019). Alumni oral histories: A collaboration between the libraries and development and alumni relations at

Grinnell College. *Collaborative Librarianship*, 11(3), 157–166.

- Humphries, L. (2015). Voices from the past. *AALL Spectrum*, 19(5), 24–26.
- Hurford, A. A., & Read, M. L. (2008). Bringing the voices of communities together: The Middletown digital oral history project. *Indiana Libraries*, 27(2), 26–29.
- Levin, T., & Doel, R. (2000). The Lamont-Doherty earth observatory oral history project: A preliminary report. *Earth Sciences History*, 19(1), 26-32.
- McKay, N. (2016). *Curating oral histories: From interview to archive*. Routledge.
- Myers, G., Lampropoulou, S. (2015). Laughter, non-seriousness and transitions in social research interview transcripts. *Qualitative research*, 16(1), 78-94.
- Powers, W. R. (2005). *Transcription techniques for the spoken word*. AltaMira.
- Ritchie, D. A. (2003). *Doing oral history: A practical guide*. Oxford.
- Sears, D. W. G. (2012). Oral histories in meteoritics and planetary science: A commentary. *Meteoritics & Planetary Science*, 47(3), 414-415.
- Weiner, C. (1988). Oral history of science: A mushrooming cloud? *The Journal of American History*, 75(2), 548-559.
- Wright, L. (2011). Macon memories: An oral history project at Randolph-Macon College. *College & Undergraduate Libraries*, 18(1), 16–36.
- Yaco, S. (2007). The potential for use of voice recognition software in appraisal and transcription of oral history tapes. *ARSC Journal*, 38(2), 214-225.