**From seeing to hearing: Lessons learned from a research-based design project focused upon audio description, print-to-acoustic remediation, and mobile-app delivery**

Brett Oppegaard

University of Hawai’i at Manoa
2550 Campus Road
Honolulu, HI 96822
+1 808-956-3340

brett.oppegaard@hawaii.edu

Thomas Conway

University of Hawai’i at Manoa
1776 University Avenue
Honolulu, HI 96822

+1 808-375-7092

tconway@hawaii.edu

Megan Conway

University of Hawai’i at Manoa
1776 University Avenue
Honolulu, HI 96822
+1 808-956-6166

mconway@hawaii.edu

**ABSTRACT**This paper describes the process we undertook to address a significant technical communication issue embedded within traditional information-distribution procedures in the U.S. National Park Service: The bureau’s nationally circulated brochures, as print artifacts, are inaccessible to people who are blind, visually impaired and print dyslexic and also not optimized for people who prefer audio communication. This report includes findings from the second year of a three-year project, funded by a grant from the National Park Service. Our first year on this project was spent getting to know the frameworks and foundations of the 342 brochures currently circulating in the country, through a deep reading of samples and a larger content analysis of the brochure structures and content genres. The second year of this project, which is the focus of this paper and poster presentation at SIGDOC 2016, revolves around the design-based research prototyping experiments we conducted as a way to create and refine a web tool to support and produce audio description, particularly in relation to the remediation of NPS brochures to audio-based mobile apps. This process led to prototyping experiments within the NPS that revealed a significant gap between practitioner knowledge and best practices. Those best practices also were found to be lacking in many areas, such as guidance for the practitioners to audio describe a NPS staple, the site map. While audio description, like open/closed captioning, might be considered a relatively straightforward process to those uninitiated in the activity, these practitioner experiments revealed many of the nuances and complexities of audio description that make it a difficult translation exercise. They also led to the generation of important scholarly research questions, such as “In what ways could one audio describe a map?”

**Categories and Subject Descriptors**

H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces – Evaluation/Methodology.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org. *SIGDOC Conference’16*, September 21–23, 2016, Arlington, VA, USA.

© 2016 ACM 1-58113-000-0/00/0010 …$15.00.

DOI: http://dx.doi.org/10.1145/12345.67890

**General Terms**

Design, Mobile, Accessibility, Usability, User Experience (UX), Audio Description, Human Factors.

**Keywords**

Universal Design, mobile, information design, cross-modal translation, National Park Service, disability studies.

# A REMEDIATION INITIATIVE

The U.S. National Park Service is the caretaker of many of the country’s most revered places. At most of these sites, a primary form of communication about the importance of the place – including foundational knowledge about its context and distinguishing features – comes in the form of a standardized paper brochure. This two-sided single-page brochure typically includes a collage of texts, photographs, maps, illustrations, and charts. Most park sites do not have a prepared audio-only equivalent for these brochures to offer if, for example, a blind visitor requests one, so various ad-hoc measures are employed, with different responses at different sites, when such a request is made. The generalized position in the bureau against more resources being allocated toward a full Universal Design of the brochures is that “not many” blind or visually impaired people visit National Parks; therefore, this issue remains a low priority. The counter argument, though, is that – according to the 2014 National Health Interview Survey, circulated by the American Foundation for the Blind [2016] – vision loss is experienced by 22.5 million adult Americans (18 and older); this vision loss affects those people directly as well as their families and friends, who all might want to visit national parks together. And they would, this argument goes, if the parks were more accessible. To test such a hypothesis, though, we needed to get more audio-described materials into park sites, to study if they indeed could create more connections with communities of visually impaired people. To get more audio description into parks, we needed to create more of them. To create more, we needed park sites to help us with the print-to-audio translation, and we needed a process through which we could work, including a specific plan, a tool for the job, and specific objectives. The imperative to test the core argument – if visually impaired people would use national parks more, if they were supported more – therefore prompted the development of this research initiative, dubbed The UniD Project.

# THE UNID PROJECT

As media in modern society generally is becoming more visual, and more oriented toward sightseeing audiences, people focused upon understanding non-visual forms of communication have sensed fresh research opportunities. One of those manifestations has been a resurging interest in audio description, including the potential for new ways to approach it academically. It has potential to be applied to contexts outside of its traditional frameworks, such as film and television. That includes toward static print media sources, as our project has done [Holsanova 2016; Maszerowska, Matamala, and Orero 2016; Matamala 2016; Matamala and Orero 2016; Ripoll and Remael 2015]. As such, this UniD project builds from previous research in audio description, yet because that research primarily has been rooted in film and television translation, this new project also has needed to reach in other directions into new disciplines and genres. In this second phase of the project’s development, it has been inspired by the historical ideas of Massimo Vignelli, which will be outlined briefly. From there, the discourse will describe how the research team has created digital tools, procedures, and protocols to advance understandings about the audio description of static media objects, such as National Park Service brochures. As part of this phase, the digital tool was tested on real users in real situations at multiple parks, which generated findings that led to iterative development of the tool and the ideas around this approach. This project is the first to attempt to audio describe NPS brochures, and make those generally accessible to people who are blind, visually impaired, print dyslexic, and audio-oriented learners and thinkers. But it also has been intended to create the groundwork (and tools and training) for a more global approach to visual media, in which everything eye-oriented in the world also could have a complementary channel for the ear to enjoy as well.

## Background and Context

In the late 1970s, designer Massimo Vignelli worked with Harpers Ferry Center staff to create the "Unigrid System," upon which all National Park Service brochures since have been based. The self-described "information architect," who also designed the innovative New York subway map, favored a modular system with a subtextual grid that facilitated order and consistency [National Park Service 2016; Vignelli 2016]. Our web-based project – with direct connections to Harpers Ferry, the National Park Service, those brochures, and those basic beliefs – has been called “UniD,” in tribute. Following that line of thinking, this project name should be pronounced like "unity," serving as both an abbreviation of the more wonky original label of "unidescription" and as an inspiration for our mission: To bring unity to the world of audio description.

Audio description (also called verbal description) can be thought of as a medium equivalent to open and closed captioning, only for audiences that prefer information in acoustic rather than visual forms. In some cases, that involves the simple verbalization of a transcript (as in text-to-voice translation), but what we mostly are concerned with in this project is the more complex audiovisual translation of visual into audible material. For example, how would you describe an Ansel Adams photograph of a scene within Yellowstone National Park to a person who cannot see, or has low vision, or has difficulty interpreting print materials, or simply prefers information in audible forms? From our perspective, those varied audiences deserve full access to this sort of public discourse, and this project has been created to serve them, under the core principles of Universal Design.

In turn, this UniD project has been developed to help people create more audio description and to serve as a robust resource for those interested in this topic. Along those lines, we have included the gathering and vetting of "best practices" guidelines as within the scope of the work, seeking the most updated scholarly research. We also are building an online forum for related thoughts and discussions. Our hope is that like the impact Vignelli's system had on NPS brochures, the UniD Project will bring higher clarity and quality to this acoustic communication form, especially in public spaces.

## Objectives of this project

Our collaborative design team – including the authors listed here, web development company personnel, and National Park Service staff members, including Media Accessibility Coordinator Michele Hartley – have spent about two years to date thinking about, discussing, and formulating objectives for this project to address compelling research questions and gaps in general scholarly knowledge about this topic but also to align with altruistic goals of improving the audio description community worldwide. Those goals included the design and development of a web tool that could do the following:

* Help users translate static visual media of any kind (texts, photographs, paintings, posters, statues, etc.) into audio content that could be freely shared, by converting text to speech, and putting those audio files directly into the hands of creators and audiences.
* Help creators manage multiple text-to-speech projects, which can be designed around any type of audiovisual translation context, including serving the need to translate a static media source (such as a brochure) or a grouping of artifacts (by theme or location).
* Provide templates for common audiovisual translation contexts and situations.
* Provide best practices and scholarly research resources related to these audiovisual translation issues and situations, helping the audio description community grow through informed, empirically based, and recursively reflective discourse.
* Provide a forum for discourse about audiovisual translation, to serve as a hub of dialectical conversations about this topic.
* Support creation of audiovisual translation that can be exported as text, audio files, and mobile apps (in Android and iOS formats), to support the broadest audiences possible.

## Method

The research team followed traditional action-research circuits [Herr and Anderson 2005] of planning, action, observation, and reflection to create multiple versions of both the alpha and beta prototype web tools (the most recent version of the development site can be examined and tested here: <http://unidescription.montanab.com/>). Park staff from three sites (Washington Monument, Golden Gate National Recreation Area, and Hawai‘i Volcanoes National Park) worked with the tool by creating audio description of their current park brochures. These prototyping exercises followed the same procedures, yet each park had an alternative approach to the tests: Washington Monument had three different staff members independently create three different audio-described versions of the brochure, to allow comparisons between authors. Golden Gate already had an existing audio-described brochure, created by a third-party vendor, and an experienced staff member, focused on accessibility issues, who single-handedly crafted the new audio-described version through our research tool. And, Hawai‘i Volcanoes decided to make the process a collective effort, in which one staff member led the process but received input from several people on site. We only prepped the participants on the use of the tool, and guided that use, so the baseline understanding and commitment to audio description at each park varied widely. The accompanying SIGDOC poster shares the results of those diverse tests in detail, including qualitative data illustrating how the different approaches generated distinct deliverables. But some of the highlights are outlined here.

## Reflection

In this process, we learned that the technical challenges of designing and developing a web tool to support better audio description were not nearly as daunting as the intellectual puzzles that emerged as we entered into this scholarly fray and found little academic literature about it and even less in terms of empirical studies that tried to make systematic sense of it.

While scholarly inquiries into other related arenas – such as open/closed captioning [Zdenek 2015], disability studies [Meena 2015], remediation studies [Zelenkauskaite 2016], mediatization [Hepp, Hjarvard, and Lundby 2015], mobilities studies [Farman 2015], etc. – provide potential models for audio description researchers to adopt and adapt, some of the issues we have faced with audio description so far have no clear and easy comparables. For example, when one entire side of a brochure is a map of an enormous park that covers hundreds of acres – including dozens of place names and visual references to locations, marked by squiggly lines, shaded sections, and complex icons, requiring a code-breaking key – the National Park Service does not differentiate whether that map is intended as navigational tool (as a way to get around) or as a piece of interpretive media (to get a sense of place). Either option seems like a legitimate use of the brochure; yet describing the ways in which to move around a place is a much different activity than describing the essence of a place and its place names, as a reference to its inscribed context. Our test subjects grappled with such complexities, and our poster will show what happened during that process, how and why.

In summary, though, there are several significant takeaways that can be offered from this research for both academics and practitioners.

* **Baseline starting points matter** – In the comparison of the three parks, one had relatively high audio-description literacy and one had very low literacy in this field (with the other one in the middle of those two). We assumed that park staff in this large system would have similar levels of training and engagement in accessibility issues, but in reality, cultural differences among the parks created a significant gap between the foundational and functional understandings of the task and projections of its final forms. For example, one park focused entirely on transcribing the text on the brochure but did not see a significant need for creating a description of the brochure’s photographs, other than to transcribe the caption text into audio form. On the other end of the spectrum, the park with a well-developed history of audio description even worked carefully on the text’s phonetic properties, so when it was machine-read, it would make more sense and sound more like natural human speech.
* **Interpretive focus matters** – One key iterative design development that came from this process was the idea that audio describers first should isolate and focus exactly on what they are describing, as a component of the overall communication content. So we created a new function in the web tool that allowed audio describers to crop an image to just what they were describing as a way to focus attention on that one piece of the larger communicative ecology. In many cases, the description focus was simple, as a separate box with text and a photo, for example. But in other cases, the brochure offered a collage with fuzzy borders and unrelated items intermingling in shared spaces, creating a noise for the interpreters that was not easily pierced. By giving the interpreters an option for visually isolating a component before trying to describe it, the early returns indicated description that also was more focused. We could compare this phenomenon to looking through a pinhole or at a full and open panorama. The pinhole view, if managed properly, could allow a clarity and purpose to the description that might get obscured in the panoramic view.
* **Subjectivity should be acknowledged and addressed** – One of the traditions of audio description is to strictly offer only the “objective” perspective in the translation. Yet, we found, as soon as a translator starts to choose what comes first, what is mentioned and what isn’t mentioned, what level of detail should be included in the description, etc., the subjectivity of the process dramatically influences the final deliverable. We also found that some of the interpreters favored more descriptive and story-based anecdotes, while others stuck tightly to the facts of the brochure, leaving little room for flourish. In that sense, this could be considered a philosophical divide, and we hypothesize that audience members likely would gravitate toward one type of description over another, probably in similar numbers. Yet we could not find any evidence-based studies that examined and tested such a fundamental concern in the field. If audio describers do not have an empirically grounded foundation for that part of the process, there is, as the other researchers cited here noted, still much work to be done.

## Conclusion

This SIGDOC 2016 paper and poster presentation might operate a bit like the middle act in a three-act play; it brought us along past the orientation and presentation of problems, through the backstory and into the thick of the mist. It showed us some paths to follow, and now, this research team and others need to go down them farther and see where they lead. In that regard, we recommend the following next steps:

* **Get more people (and parks) to try audio describing their materials** – Small and anecdotal samples can raise many compelling questions, so we could benefit as a field, even if it eventually gets subsumed as a subfield of some other discipline, by getting more people to do this kind of work and raise their ideas and questions about it.
* **Seek not just questions but also answers –**When researchers uncover these sorts of questions – and many of them have been uprooted so far in this project already – we need to start pinning them down as much as possible if this field is going to mature and develop beyond its mostly ad-hoc traditions and anecdotal approaches to the translation craft. Like an apprenticeship, audio describers often learn the process from those who have come before them. But mostly missing in that dynamic to date have been the people (scholarly researchers) to challenge those traditions and force them to survive on evidence of their effectiveness rather than the notion of “this is the way it’s always been done,” or “this is the way my mentor did it.” In short, audio description needs more people to try to establish through empirical evidence that the traditions have validity beyond cultural grooming.
* **Audio description has potential to break out of the box** – While this field was founded in description of television, film, and stage events, it also has significant potential for the rest of the visual media world, including for text, print artifacts, and photographs. If you think of those as forms of visual media, inaccessible to those who cannot see, and then spend some time reflecting upon and thinking about all of the dark spots in a community’s discourse that could be lit up by audio description, you can begin to imagine the potential of this project, and its freely accessible tools, and support systems, designed to allow the audio description of anything anywhere.

This project so far has been mostly about recognizing the potential of audio description, in the National Park Service and beyond. It has been about creating the groundwork for future studies. This second phase of the project has started to strip away the assumptions and rote processes intertwined in the field’s infancy. But it also has highlighted just how much work needs to be done and explored at least one way of how to do it.

Our approach at this point has been open and formative in nature, but as we dig deeper into wayfinding around these related issues, the more specific scholarly paths of inquiry have emerged. Within about the next year, we plan to build and develop audio-described NPS brochures at about 40 sites across the nation, greatly increasing our sample size, types of data collected and overall amounts of data collected. By using an action-research approach, and developing our understandings iteratively, round by round, we hope to tease out the academic areas in the most need of attention, isolate them for study, and then seek out answers that will help the audio description field, audio describers, and better understand audio-visual translation in these contexts but also in general. Vignelli’s design vision compels us to look for straightforward solutions, elegant in their simplicity. But getting to the end of the process, like the end of the subway line – in terms of translating the print brochures into audio forms – is only a part of this project. It’s the in-between stops, like this one, that are the most interesting, as they give us a chance to carefully think about the audio description process and products to determine what we can see – and hear – when more people get involved in this type of translation activity.

# REFERENCES

1. American Foundation for the Blind. (2016). Retrieved April 15, 2016, from <http://www.afb.org/info/blindness-statistics/adults/facts-and-figures/235#demographics>.
2. National Park Service. (2016). Retrieved April 15, 2016, from<https://www.nps.gov/parkhistory/online_books/brochures/unigrid/index.htm>.
3. Massimo Vignelli. (2016). Retrieved April 15, 2016, from <http://www.vignelli.com/awards/massimo.html>.
4. Zdenek, S. (2015). *Reading Sounds: Closed-captioned Media and Popular Culture*: University of Chicago Press.
5. Meena, S. (2015). Disability studies and scope for rehabilitation of differently abled children. *Journal of Disability Studies, 1*(1), 35-40.
6. Zelenkauskaite, A. (2016). Remediation, convergence, and big data Conceptual limits of cross-platform social media. *Convergence: The International Journal of Research into New Media Technologies*.
7. Hepp, A., Hjarvard, S., & Lundby, K. (2015). Mediatization: theorizing the interplay between media, culture and society. *Media, Culture & Society*,.
8. Farman, J. (2015). *Foundations of Mobile Media Studies: Essential Texts on the Formation of a Field*: Routledge.
9. Matamala, A., & Orero, P. (2016). *Researching Audio Description: New Approaches*: Springer.
10. Matamala, A. (2016). The ALST Project: Technologies for Audio Description *Researching Audio Description* (pp. 269-284): Springer.
11. Holsanova, J. (2016). Cognitive approach to audio description. *Researching audio description: New approaches*, 49-73.
12. Ripoll, A. M., & Remael, A. (2015). Audio-description reloaded. *Translation Studies, 8*(1), 63-81.
13. Maszerowska, A., Matamala, A., & Orero, P. (2014). *Audio Description: New Perspectives Illustrated* (Vol. 112): John Benjamins Publishing Company.
14. Herr, K., & Anderson, G. L. (2005). *The action research dissertation: A guide for students and faculty*. Thousand Oaks, CA: Sage Publications, Inc.