

# Centering Revitalization in Remote Documentation

Meg Harvey\*

*University of Arizona*

## Introduction

While there are clear benefits to language documentation in endangered language communities, when it is not done thoughtfully and with community goals at its core, it can take away valuable time that could be spent on work that more directly creates new speakers. Furthermore, documentation made from materials intended for linguistic analysis and description can be obtuse and difficult to adapt to pedagogical materials (Nathan & Feng 2009; Langley et al. 2018). Previous research (e.g., Yamada 2011; Fitzgerald & Hinson 2013) avoids this pitfall in face-to-face documentation by structuring it as a feedback loop, with capacity building, documentation, and analysis each feeding into the other. Silva and Dupris (2020) expand on this by ensuring that revitalization output and activities (like teaching materials) are the core goal of documentation work and that revitalization is returned to at each step in the feedback loop. In doing so, they ensure that revitalization and community goals are centered throughout the process.

I extend this revitalization-centered methodology into remote fieldwork by laying out a method for documenting languages online that a) centers revitalization activities and goals, b) limits the work required on the individual participants' end, and c) produces clear recordings that can be used for revitalization goals and linguistic analysis. Specifically, I discuss my experiences documenting the Tunica language online. Section 1 briefly introduces the Tunica language and current efforts to revitalize the language. Section 2 delves into the specifics of the remote documentation process, including the four types of documentation efforts undertaken in this project and their outputs for revitalization and

---

\* I would like to thank the mentors and apprentices at the Tunica-Biloxi Language and Culture Revitalization Program for their guidance and advice – Tikahch! Thank you as well to Andrew Abdalian, Raina Heaton, Patricia Anderson, Judith M. Maxwell, Robert Henderson, Ryan Bennett, Benjamin Eishchens, Amy Fountain, Amber Lubera, and the anonymous reviewers for their valuable comments.

analysis. Finally, ‘Discussion’ frames this within the larger context of remote documentation of endangered languages.

### **1 The Tunica Language**

Tunica is a reawakening isolate spoken by the Tunica-Biloxi Tribe of central Louisiana (Anderson & Maxwell 2019: 45). A reawakening language is a language whose transmission was interrupted; meaning that new speakers are learning from documentation of fluent speakers contained in documents like grammars, recordings, and texts rather than directly from those fluent speakers. There are currently at least 32 second-language speakers of Tunica, both on the reservation in Marksville, Louisiana, and throughout the United States (Eberhardt et al., 2020). Every current speaker of Tunica has learned the language either directly from documentation of fluent speakers or, more frequently, from language lessons adapted from these texts.

The Tunica-Biloxi Language and Culture Revitalization Program (LCRP) serves the Tunica-Biloxi community as the main organization driving revitalization efforts. While the eventual aim of the program is to get Tunica spoken daily in many different domains, the current focus is on the development of a Mentor and Apprentice program wherein three language mentors train 5 language apprentices over the course of three years. The rest of the revitalization program focuses on immersion learning in a classroom setting. There are three regular language classes: one aimed at children aged 5-10, one at children aged 11-16, and one at adults aged 17 and older. These classes are supported by an introductory textbook which is intended to be used across age ranges and by teachers, students, and solo learners alike (KYLY m.s.).

### **2 Remote Documentation Methods**

For this paper, I recorded eight participants in the LCRP, three of whom were language mentors (including myself) and five of whom were language apprentices. Although this is a small number of participants, it constitutes everyone who regularly teaches Tunica language classes.

The program’s focus on classroom education means the documentation process must put minimal stress on teachers and students and, ideally, result in new pedagogical materials and observations that can help teachers refine their existing materials.

Another aim of this process is to provide self-study materials for learners. While much of the LCRP’s programming has moved online since 2020, the time of the courses

and activities can still limit the participation of working adults and others on a strict schedule. The textbook can provide some support for these learners but is particularly limited when it comes to modeling pronunciation and extended conversation. These two gaps are ones that these remote documentation methods are well-positioned to address.

From the linguistic analysis side, there are some very real concerns about the generalizability of audio data collected online: software, equipment, and an internet connection can all affect the quality of an audio recording, even rendering a recording largely useless for phonetic data (Sanker et al. m.s.). In keeping with the recommendations of Sanker et al., I ensured that everyone was working with the same equipment. To this end, all participants used Cooler Master MH630 headsets, which have a sampling rate of 44.1 kHz (recommended for phonetic documentation) (Thauvette et al. 2020; Sanker et al. m.s.). All participants also used DuKabel USB to 3.5mm Jack Audio Adapters to hook the headsets into the USB port instead of the audio jack.

The recordings themselves were made using the website Zencastr ([www.zencastr.com](http://www.zencastr.com)). One advantage of Zencastr is its ease of use: participants only needed to wear their headsets and join the recording session in a supported browser; I handled the rest. Because documentation during online classes required that they be on Zoom simultaneously, some participants did need to be trained in lowering the volume on specific programs to avoid echoes. Ultimately, Zencastr is a relatively stable, user-friendly resource for online data collection, provided participants have consistent internet access.

While this website is primarily concerned with podcast recording rather than language documentation, its premium edition allows the host to collect individual .wav files from each participant. I then used the audio editing software Audacity to remove identifying information (e.g., names or private discussions held in class) and then annotate each resulting individual file in Praat (Audacity Team 2021; Boersma & Weenink 2021). Additionally, I combined the edited .wav files from each session and created time-aligned transcripts in both Tunica and English using ELAN (ELAN). In sections 2.1 - 2.3 below, I discuss how this process and the outputs varied with each documentation method and how each method supported the larger revitalization goals of the community.

## **2.1 Classroom Documentation**

The bulk of the documentation in this project comes from documenting the online Apprentice Language Classes. These are teacher-training courses conducted primarily in Tunica with the goal of preparing the apprentices to teach lessons and manage the

classroom without using any language other than Tunica. The table below summarizes the revitalization output and analytical utility of the files created during classroom documentation.

Table 1. Classroom Documentation Summary

<b>File Type</b>	<b>Revitalization Output</b>	<b>Use for Analysis</b>
Mono file with all participant's recordings with ELAN transcription	Audio and time aligned transcript of sentences and words in spoken Tunica and a transcript of the translations.  Can include examples of certain constructions (like imperatives and questions) that rarely show up in other elicitation styles, such as extended personal narratives.	This is effective for analyzing phonology, phonetics, morphology, and syntax in short, spontaneous speech. Particularly effective for looking at question formation.
Individual .wav files with PRAAT transcription	Audio of sentences and words in spoken Tunica (largely unconnected.) Potentially useful as part of a talking dictionary or to pair with the textbook.	Particularly effective for analyzing phonetics and phonology and looking at individual variations.

## 2.2 Conversation

“Conversation” describes a broader elicitation type, in which participants have a short discussion concerning a pre-provided topic. There are not many opportunities for extended spontaneous speech in Tunica, particularly with such a large group. To account for any anxiety speakers may have about being recorded, participants are provided with the core topic of conversation ahead of time. Allowing people to prepare is the best way to get extended speech, but I note which parts of the conversation were prepared rather than

impromptu. In addition to the documentary and linguistic benefits of collecting this information, these interviews double as a podcast episode for language learners. Because many Tunica community members are living away from the main hub of language revitalization efforts in Louisiana, these podcasts constitute a vital chance for non-local learners to be exposed to varied, extended conversations in the language.

Table 2. Conversation Summary

File Type	Revitalization Output	Use for Analysis
Mono file with all participant's recordings with ELAN transcription	A podcast format conversation concerning different topics. It allows listeners to hear extended, connected speech in Tunica.	Effective for analyzing morphology and syntax in short, largely spontaneous speech. Potentially valuable for semantics and pragmatics as well.
Individual .wav files with PRAAT transcription	Audio of sentences and words in spoken Tunica. Potentially useful as part of a talking dictionary or to pair with the textbook, however "word lists" may be better for this.	Particularly effective for analyzing phonetics and phonology and examining individual variations in speech.

### 2.3 Translation Tasks

Unlike the classroom documentation and conversation methods which included both a core revitalization activity and a revitalization output, the translation task takes a more traditional approach to elicitation. The translation tasks ask participants to take sentences or words in English, translate them into Tunica, and read the Tunica translation aloud. This task will target certain sentence constructions. For example, certain question formation patterns vary amongst speakers. (1) below shows two possibilities for the question "How many years?".

(1)

- (a) *Tihika kashku?*  
 tihika kashku  
 year[PL] how.many  
 'How many years?'

- (b) *Kashku tihika?*  
 kashku            tihika  
 how.many        year[PL]  
 ‘How many years?’

Example 1a shows the form taught in the textbook, wherein *kashku* follows the noun phrase. Example 1b, however, shows a form frequently taught in the classes, wherein speakers place the question word *kashku* before the noun phrase. This construction is influx and understanding how people are using it will help the LCRP tailor its material to reflect the language as its being used.

Table 3. Translation Task Summary

File Type	Revitalization Output	Use for Analysis
Individual .wav files with PRAAT transcription and ELAN file	Audio of sentences and words in spoken Tunica. Useful for creating examples of sentence structures that learners may want to practice or as part of a talking dictionary. Additionally serves to ensure that language materials reflect what speakers are actually saying.	Particularly effective for analyzing syntax and morphology.

## 2.4 Word Lists

Like the translation tasks, word lists do not center around an existing revitalization activity but still yield valuable results for revitalization and analysis. One focus of word list tasks was to target difficult to identify . For example, Tunica as it is currently taught has a phonemic pre-aspirated voiceless velar stop ([<sup>h</sup>k]) that has proven difficult for learners to identify and produce. Word lists provide excellent perceptual practice.

Table 4. Word List Summary

File Type	Revitalization Output	Use for Analysis
Individual .wav files with PRAAT transcription and ELAN file	Audio of individual words spoken in Tunica. Useful in a talking dictionary or even structured into lessons for learners practicing certain sounds (ex: [ʰk])	Particularly effective for analyzing phonetics and phonology at the word-level and looking at individual speaker variation on certain sounds.

### Discussion

This paper describes a method for remote revitalization-centered documentation of reawakening languages. For this type of work to be most effective, it either must integrate with existing revitalization activities (as the ‘Classroom Documentation’ method does) or create a new opportunity for group language use (like the ‘Conversation’ method). These core approaches can then be supplemented by minimally-taxing elicitation techniques like word lists and translation tasks, which allow language workers and researchers to target specific revitalization or documentary output with minimal stress for speakers. These recordings can then be used for documentation, revitalization, and analytical purposes, provided the recordings are made using consistent, good quality software and hardware.

By embracing the fact that Tunica language activities had moved online, this project reduced the onus on the language worker participants and produced new content for speakers. While these methods are not appropriate for every language community or type of analysis, they form a strong foundation for linguists doing fieldwork online. Indeed, researchers and language workers may consider including this type of documentation in their repertoire even when in-person documentation becomes available.

### References

Audacity Team (2021). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 3.0.0 retrieved March 17th 2021 from <https://audacityteam.org/>

- Anderson, P. M. & Maxwell, J.M. (2019). The Tunica Language. *Language in Louisiana: Community and Culture*, ed. by Natalie Dajko and Shana Walton, 45-63. Jackson: University Press of Mississippi.
- Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.2.03, retrieved 3 December 2021 from <http://www.praat.org/>
- Eberhard, D.M., Simons, G.F., & Fennig, C.D. (eds.). 2020. *Ethnologue: Languages of the World*. Twenty-third edition. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- ELAN (Version 6.2) [Computer software]. (2021). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from <https://archive.mpi.nl/tla/elan>
- Fitzgerald, C. M. & Hinson, J.D. (2013). *Foundation for Endangered Languages, FEL XVII: Endangered Languages Beyond Boundaries: Community Connections, Collaborative Approaches, and Cross-Disciplinary Research*, 2013. The Foundation for Endangered Languages: Bath, England; 2013.
- Kuhpani Yoyani Luhchi Yoroni (KYLY). (m.s.). *Rowina Taworu Luhchi Yoroni (Tunica Language Textbook)*. University of Indiana Press. Bloomington, Indiana.
- Langley, B., Langley, L., Martin, J., & Hasselbacher, S. (2018). *The Koasati Language Project: A collaborative, community-based language documentation and revitalization model: From Theory To Practice Around The Globe*. 10.1515/9783110527018-007.
- Nathan, D. & Feng, M. (2009). Language documentation and pedagogy for endangered languages: a mutual revitalisation. *Language Documentation and Description*, vol VI, 132-160.
- Sanker, C., Babinski, S., Burns, R., Evans, M., Kim, J., Smith, S., Weber, N., & Bower, C. (2020). (Don't) try this at home! The effects of recording devices and software on phonetic analysis. Unpublished manuscript. Retrieved from <https://lingbuzz.net/lingbuzz/005748>
- Silva, W., & Dupris, J. (2020). On a method for language documentation based on language (re)vitalization activities: a case study. *Cadernos de Linguística*, v. 1, n. 3, p. 01-14, 17 Dec. 2020.
- Thauvette, J., McCaughey, V., & Khong, Y. (May 27, 2020). Cooler Master MH630 Headphones Review. [rtings.com](https://www.rtings.com)

<https://www.rtings.com/headphones/reviews/cooler-master/mh630#page-test-results>

Yamada, R. (2011) Integrating Documentation and Formal Teaching of Kari'nja: Documentary Materials as Pedagogical Materials. *Language Documentation & Conservation* 5. 1-30.

Zencastr, Inc., January 2021. Zencastr. [www.zencastr.com](http://www.zencastr.com)

© Meg Harvey

*The University of Arizona*

mharvey3@email.arizona.edu

*Coyote Papers Volume 24 (2022)*