Reflections on (de)colonialism in language documentation

Wesley Y. Leonard
University of California, Riverside

With origins in colonial logics and institutions, language documentation practices can reinforce colonial power hierarchies and norms in ways that work against the needs and values of Indigenous language communities. This paper highlights major patterns through which this occurs, along with their effects, and models how language documentation can be structured in ways that are more grounded in the experiences and perspectives of the communities that use it. I propose decolonial interventions that emerge from Indigenous research principles and perspectives, and illustrate how these practices can better support language community needs while also improving the scientific value of language documentation.

1. Stories of engagement with language documentation

As a linguist who focuses on reversing language shift in Native American communities, much of my professional work involves promoting language documentation, using its resulting products, and engaging with the broader social issues that surround it. For example, I have served many times as an instructor for Breath of Life programs, in which Native Americans access and interpret archival documentation for language reclamation purposes. These programs include training to facilitate the use of legacy documentation, much of which was created by linguists and almost none of which was created for pedagogical purposes. This process of consulting language documentation to reverse language shift is very close to me since it is what occurred in my own Miami community: Our tribal language, myaamia, was sleeping for about 30 years and was later brought back into community use from archival records (Leonard 2008).

I would like to thank Colleen Fitzgerald, Andrea Berez-Kroeker, and an anonymous reviewer for their feedback on an earlier version of this paper. I also offer tremendous gratitude to the participants of the Natives4Linguistics satellite workshop at the 2018 Linguistic Society of America Annual Meeting for sharing their perspectives, which are embedded throughout this paper. Funding for the Natives4Linguistics workshop was provided by the National Science Foundation, BCS grant #1743743.

As I also hold a disciplinary interest in language documentation, I share other linguists’ concerns about documentation methods, archiving, and the like, and welcome the growing body of literature on these topics with its growing focus on language communities. However, while language community members are increasingly written about, they still rarely serve as primary voices in scholarly outlets. I thus focus this paper on my perspectives as an Indigenous community member, specifically a citizen of the Miami Tribe of Oklahoma, who gained access to myaamia because of documentation. This is a story grounded in lived experiences—my own, and those that have been shared with me by other Indigenous people.

Common to these stories is an emphasis that the (non-)use of a given language emerges from social marginalization that Indigenous people continue to experience. My tribal community shifted away from myaamia almost entirely, and while there are multiple specific causes, the basic underlying theme is colonialism. By this, I refer to ideas and practices of subjugation by socio-politically dominant groups or institutions (including academic disciplines) that assert and maintain control over the minds, bodies, and cultures of other groups, generally with an intent of exploiting them to benefit the dominant group. Decolonialism, by extension, disrupts the ideas and institutions of colonialism. It is a way of thinking and acting that emphasizes the sovereignty, peoplehood, intellectual traditions, and cultural values of groups that experience colonialism. As with the language-specific decolonial movement that I call language reclamation, which refers to revitalization efforts that are grounded in and driven by community needs and values (Leonard 2011, 2012, 2017), decolonial approaches in general look not just at a given current situation (e.g., “Language X has only five speakers so we must document it now”) but also at the histories of institutions and power that have fostered it (e.g., “Why does Language X have only five speakers—and whose definition of ‘speaker’ is being used? What broader issues occurred in Community X to create this situation?”)

As communities are diverse, decolonial interventions must be specific. The ideas in this essay are primarily informed by experiences of North American Indigenous communities and are offered as examples. Colonialism, however, while also realized with respect to specific places, peoples, and contexts, is manifested in Documentary Linguistics in more general ways. This is because Documentary Linguistics emerges largely from a Euroamerican colonial tradition that has guided the development of Linguistics (Errington 2008), whose scope is global but whose actors are concentrated in institutions that follow Western traditions of research. These traditions establish languages as objects to be described in scientific materials (e.g., texts, corpora, technical publications) which can serve multiple audiences, but normally are structured around colonial categories and norms of description. Examples include how languages are classified (e.g., with numerical vitality scales), named (and given ISO 639 codes), and written. Lise Dobrin and Josh Berson (2011:202) capture this pattern well in their critical analysis of language documentation, noting how “linguists’ scientific authority … takes for granted one group’s power, derived from its association with the high-status western institution of the academy, to cast its gaze upon cultural others through the

---

3 Terms such as (de)colonialism, (de)coloniality, (de)colonization, settler colonialism, and related concepts such as imperialism are used in different ways, the details of which go beyond the scope of this paper. I recommend that people engaged in language documentation focus more on the underlying ideas rather than the terms.

4 I adopt the convention of capitalizing disciplinary names, but using lower-case to refer to the associated research.
research process, and to represent them according to its own, externally imposed analytic categories in the resulting scholarly products.”

The hegemony of academic fields is so strong that they are easily assumed to be the logical unit of analysis, and thus the default starting point from which to develop theory and research questions. For investigating (de)colonialism in Documentary Linguistics specifically, one might, for example, focus on a person whose work has had a large influence on language documentation (e.g., John Peabody Harrington), on a movement (e.g., salvage linguistics), or on a seminal publication (e.g., Himmelmann 1998). However, while potentially illuminating, a problem with this approach is that it can reproduce colonial hierarchies by elevating named academic fields over the much broader sets of lived experiences and issues that underlie language documentation needs. I thus instead draw attention to the experiences of people whose languages are the focus of recent documentation efforts.

Members of these language communities in many cases engage with Documentary Linguistics, particularly via the prototypical model whereby trained linguists from outside the community, most of whom have advanced academic credentials, work with speakers of the language under consideration to intentionally create records of it primarily for scientific purposes and secondarily to support other needs, such as language teaching. I am very familiar with this approach, but through my Miami lens, hearing the term ‘language documentation’ immediately raises other topics: The forced removal in 1846 of my direct ancestors from tribal homelands in Indiana is part of the story. The underlying intent of the Jesuit missionaries who created the first written myaamia records comes to mind as well. My ancestors’ experiences in Indian boarding schools are part of the story. My own experiences in educational institutions as a scholar who regularly has to explain basic tenets of Indigeneity, such as the fact that Native Americans still exist, are likewise part of the narrative. Although I have not personally done much direct analysis of myaamia documentation, I have heard anecdotes from others who have had trouble with it because of how myaamia is represented. Through engagement with these and similar stories, Documentary Linguistics can become more decolonial. Below, I provide a critical examination of language documentation as a colonial enterprise and offer suggestions on how to move it toward a decolonial practice.

2. Colonial approaches to language documentation Colonial approaches are not intrinsic to language documentation, as Indigenous peoples have made efforts in this area through oral traditions and cultural practices that move their languages and the associated intellectual traditions across generations. My experience, however, is that unless Indigenous community members create a systematic record of their languages in alignment with the standards of Documentary Linguistics, these efforts are not recognized as ‘language documentation’. Defining the scope of ‘documentation’ as a category is a manifestation of power. Ironically, the development of Documentary Linguistics as a named academic field, while generally beneficial for Indigenous communities, may also serve to constrain what counts.

The previous example focuses on demarcating ‘documentation’, especially with respect to its being defined separately from other activities (e.g., description) that produce similar products and are thus often similarly experienced by members of Indigenous
communities. Even more important is how ‘language’ is defined (Leonard 2017). Contemporary linguistic science privileges certain ways of defining language, particularly by structural units that can and often are described and analyzed not only separately from each other, but that are also disembodied from the people who use them, thus contradicting Indigenous values of interrelatedness as a framework for describing and interacting with the world. This trend of conceiving of languages as structurally-defined objects emerges in linguists’ analyses of Native American languages, which Joseph Errington (2008:8) observes cover “enormously different languages, [but] also resemble each other in obvious ways ... [because] each describes an object which falls under a single, common category.”

Language documentation practices can often also impose colonial norms of analyzing language in ways that misalign with the needs and values of Indigenous communities (Grenoble 2009; Hermes, Bang, & Marin 2012; Mellow 2015; Leonard 2017). ‘Dissecting’ is the word I tend to hear in critiques from Native American community members who are working with language documentation, particularly legacy documentation, and opine that some linguists’ approach to investigating language is inappropriate or offensive. This occurs, for example, when language community members encounter a linguist (or a material created by a linguist) that presents a grammatical issue as a puzzle to be solved for ‘our understanding’ (where the pronoun seems to refer to other linguists). Solving this puzzle then occurs through language data isolated from cultural contexts, and the analysis fails to acknowledge the people who claim the language, let alone to engage with what the language represents to them.

Even when all stakeholders appreciate this documentation and the people who have contributed to creating it, there is an ongoing problem of linguist-focused materials being inaccessible or otherwise misaligned with community needs. For example, a documentation corpus may contain carefully annotated texts but lack examples of conversations. The current movement in language documentation projects toward prioritizing domains that the community considers important and broadening the scope of what gets documented (e.g., by including video of conversations on diverse topics) represents a significant improvement. However, a decolonial approach calls not just for considering community needs, but rather for starting with them in conceiving of language documentation as an idea, as well as for developing the specific methods and goals of a given documentation project. Next, I present some possible interventions for accomplishing this.

3. Decolonial approaches to language documentation Were Linguistics to make a full decolonial shift, I believe that documentation would remain very important but that the norms of planning and implementing documentation projects would be driven by Indigenous research methods and protocols, which are decolonial by design and exemplified below. Due to length limitations, I focus here on recurrent themes in Indigenous research: centering details in whole systems (e.g., crafting language documentation in reference to how a language exists in its full context, and how communities want it to exist in the future), focusing on relationships and reciprocity, respecting the responsibility that comes with knowledge and its dissemination, and

\[5\text{For the remainder of this paper, I will incorporate description, research, and products that emerge from documentation projects within my analysis of ‘language documentation’ in response to my experiences in Indigenous community settings, where these all tend to be spoken about as a single thing.}\]
actively engaging with community needs and institutions at all stages of the research process.\footnote{These ideas have been developed and described by many Indigenous scholars (e.g., Wilson 2008; Kovach 2010; Smith 2012). For a short synthesis of Indigenous approaches in science, see Bang, Marin, & Medin (2018).}

Using an Indigenous approach, a clear requirement for language documentation work is engagement with community definitions of ‘language’ and with community beliefs and analyses about how it functions. For instance, in my Miami community, as with many other Indigenous communities (e.g., Shaw 2001; Nelson 2002), there is a strong focus on our language’s relationship to land and an associated belief that language vitality requires community access to our lands. As such, documentation practices that fail to acknowledge the land serve to erase our connection to it, thus reinforcing the legacy of colonial violence that dispossessed Miami people of much of our land. Appropriate land acknowledgements represent a simple yet significant intervention.

Beyond an emphasis on land, also common in Indigenous community contexts is for language and peoplehood to be considered heavily intertwined (see, e.g., Clarke 1996; Meek 2010). From this point of view, representations of a language become representations of the people who claim it, and by extension also of their political sovereignty. As such, presenting the language as an object whose value lies in what it reveals for linguistic theory can reduce the people to their value for science, thus evoking the general colonial practice of exploiting the colonized population for its resources. Notably, most Indigenous community members I have spoken to about this do not mind the idea that their language’s structure will inform linguistic theory so long as the community’s ideas about language are respected. This, however, cannot easily happen, even by well-intentioned users of language documentation, unless the community’s ideas about language are present and prominent in it.

Emerging from how ‘language’ is understood are the norms of analyzing it, which also have significant implications. When documentation materials default to discrete structural units in presenting a language, the associated people can symbolically be reduced to discrete parts as well. While difficult to avoid in some situations, one useful practice when disseminating documentation is to feature larger and well-contextualized language examples, initially represented as the community would most commonly represent them (which of course first entails asking about this), and only after establishing this norm to represent them in other ways that may be necessary for specific tasks such as presenting a morphological analysis. This is especially true for language examples provided as interlinear glosses, where current disciplinary conventions allow for the initial line in a given example to be presented with the author’s morphological analysis—i.e., with hyphens, periods, and odd spacing. As argued by Wendat linguist Megan Lukaniec (2018), this is inappropriate; rather, it is important to start an example using unbroken words.

Beyond language definitions and analyses about its functions, ideas about the appropriate ways of using language must also be core to language documentation. In making this claim, I follow Jane Hill’s (2006) call for more ethnography in language documentation, where ideas about language will be emphasized on par with grammatical and lexical information. My observation has been that ethnography, though praised in its own right, is often talked about as something separate from ‘language documentation’
Reflections on (de)colonialism in language documentation

among linguists (except insofar as cultural patterns emerge from lexicons and texts).\(^7\)
In all of my experiences in Indigenous communities, however, this distinction has been perceived as strange, and of course a common way for community members to define ‘language’ is with reference to culture (Leonard 2017).

Fortunately, shifts are occurring in Documentary Linguistics and in academia more widely such that colonial norms and ideas are being challenged. For instance, there are a number of projects in which community members work in successful collaborations with outside researchers who are firmly committed to supporting community needs and values in creating documentation (see, e.g., Hermes & Engman 2017; Genee & Junker 2018; and the essays in Bischoff & Jany 2018). Moreover, evidence of community engagement, support, and accessibility is normally now expected by documentation funders. A related decolonial intervention occurs with reworking legacy documentation to make it more accessible to language communities, as occurs in my Miami community.\(^8\) Along with this issue of accessibility, broader ethical issues are now a topic of focus, with scholarship on these issues (e.g., Grinevald 2006; Warner et al. 2007; Rice 2010, 2011) becoming standard reading for training in Documentary Linguistics. I have been especially inspired by scholarship on improving language archives by integrating community needs and insights (e.g., Linn 2014; Shepard 2016), and by the observation that community needs and insights are fundamental to a feedback loop wherein documentation, analysis, revitalization, and training work in conjunction with each other such that each is improved relative to what it would be on its own (Fitzgerald in press).

Characterizing the examples cited above is a strong awareness of how documentation is actually used (or not used) on the ground in community contexts, and how interpersonal dynamics play into the associated outcomes. From this emerges a key principle, which is that language documentation, while presumably intended to describe how a language works, can actually be prescriptive in how it is received. That is, documentation creates a baseline with prescriptive implications because it is people who create it and use it, and people have backgrounds and power relations with each other. While this is true across the board, for purposes of moving language documentation toward a decolonial practice I believe it is most useful to highlight the prototypical situation referenced earlier in which a professional linguist, who is not a community member, is the primary person who curates documentation materials and analyzes the language. When these professionals say to members of language communities things like “this is how your language’s grammar works”, their words can easily come across as fixed truths rather than what they actually are—analyses by specific people who have specific backgrounds with respect to age, gender, ethnicity, and other traits. Other practices that can yield ‘truths’, and that thus must be performed with care, include how languages are classified with respect to vitality (e.g., Leonard 2008, 2011), how speakerhood is determined and valued (e.g., Grinevald 2003; Leonard & Haynes 2010; Dobrin & Berson 2011; Muehlmann 2012; Boltokova 2017),\(^9\) and how languages are written with respect to orthographic choices (e.g., Romaine 2002; A reviewer notes that this split also stems from the disciplinary boundaries that emerge from colonialism and are reinforced by colonial power structures. Indeed, this is true, and breaking down disciplinary silos is part of decolonial work.

\(^7\) Other examples of successful decolonial interventions with legacy linguist-oriented materials appear in Warner et al. (2006), Oberly et al. (2015), and Langley et al. (2018).

\(^8\) I have observed that language consultants are too often reduced to ‘speakers’ in linguistic science, even though they are actually full people with kinship networks, occupations, responsibilities, needs, hopes, and intellectual contributions that go beyond their linguistic knowledge.
Leonard Oko 2018) along with broader issues of transcription and entextualization (see Bucholtz 2007; Riley 2009).

The effects of these decisions are heavily intertwined with the positionalities of the scholars who engage in language documentation, particularly the relationships they have with language communities and with the academy. It is thus crucial that these scholars be reflexive about how their personal backgrounds may guide how their work is perceived, and that they practice self-location (i.e., explicitly acknowledge their positionalities) in research contexts. The same principle applies to community members, particularly with respect to understanding how the roles that individuals are expected, allowed, encouraged, or discouraged to have in documentation projects reflect and affect their other community positions. A recurring example in my experience with small communities in the United States is that ‘elder’ gets overly linked to ‘speaker’ in documentation contexts even though this intersection of roles is not traditional, but rather a circumstance of language shift. Indeed, if a given language is known only by elders, it follows that they are going to play a pivotal role in most documentation projects; this in itself I see as sensible. The problem I have observed is that this situation too easily intersects with colonial ‘dying language’ (Leonard 2008) and ‘last speaker’ (Davis 2017) discourses that disallow younger and future generations to ever be legitimate speakers of a given language because a traditional role of elders, that of respected knowledge-bearer, has evolved through conversations about language documentation to constrain ‘real’ speakerhood. Being decolonial in language work entails calling attention to this sort of issue, which cannot be resolved unless people are aware that it might occur and thoughtful about how they advocate for, implement, and disseminate the results of documentation projects.

4. Concluding thoughts My commentary about the value of decolonial practices in language documentation is based on the following belief, whose ensuing social justice aims I have observed to be increasingly commonly proclaimed in Documentary Linguistics: Colonialism is bad, and decolonial interventions are thus appropriate. I have encountered only a few scholars who state otherwise, usually under the guise of promoting ‘objective’ science, though more common (and more troubling) are the problematic statements I have heard from language documentation practitioners who claim to support social justice but whose actions suggest otherwise. These include disparaging commentary about Indigenous groups, anecdotes about skirting tribal cultural values or research protocols, and statements that linguists are the people who truly understand language. Fortunately, this type of thinking is becoming less common, and regardless, the decolonial interventions I propose really should not be controversial, even among scientists whose goals are not focused on social justice, since these proposed practices also improve language documentation by facilitating a more complete “record of the linguistic practices and traditions of a speech community” (Himmelmann 1998:166). This noted, colonialism is so engrained in the academy that abolishing it is difficult, even when doing so arguably leads to better science.

In reference to addressing this challenge, contrary to my earlier point about looking beyond academic disciplines as units of analysis, here I make a call to academic disciplines to address harmful practices through their professional structures. This can occur, for

---

10See Riddell et al. (2017) both for an example of researchers doing this in a publication, as well as for the authors’ excellent discussion of how self-location guides ethical research by facilitating an understanding of power differentials.
example, in a Linguistics department, which can foster a norm of including Indigenous ideas about language in introductory courses and of continuing this practice in training that is more specific to language documentation, such as field methods courses. Training in the technical aspects of documentation and linguistic analysis can be complemented by coursework in Indigenous research methods and ethnography. Professional organizations in language sciences can also advocate for decolonial practices and provide supporting guidance on the hiring, promotion, and funding of practitioners in the field. Closely related to this is the appropriate use of what is generally considered the gold standard of quality in academic work—peer review. When the language community is recognized as a core stakeholder, it follows that members of the language community will be among the reviewers of language documentation proposals and products.11

I will end with my reflections on the question that I am most often asked by linguists during discussions about decolonizing language documentation: “What can I do?” Several specific decolonial practices that individuals can undertake have already been addressed above, but I intentionally leave a specific discussion about individuals’ actions for the end. I do this to illustrate yet another principle of counteracting colonialism, which is the need to shift the focus away from individual approaches in isolation to instead first center the larger social norms and institutions in which individuals operate. Stated more directly, colonialism is endemic and institutionalized; therefore, counteracting it in language documentation (and beyond) involves addressing colonial structures. This means that individuals must be deliberate, ideally also explicit in their research products, about counteracting colonialism, and open about where they need support. I hope in the future that decolonial forms of language documentation will be the default and thus unnecessary to call attention to, but to get to that stage we must all be thoughtful and intentional when doing language work.

11Due to length limitations, I am not discussing yet another level of finalizing and disseminating this work: intellectual property and associated legal instruments, such as copyrights. I recommend Madsen (2008) for general discussion of this topic in Indigenous research, Tatsch (2004) for discussion about language as intellectual property, and Langley et al. (2018) for useful examples from a collaborative language documentation project.
References


Wesley Y. Leonard
wesley.leonard@ucr.edu
orcid.org/0000-0001-8792-4414