Language Documentation and Description

_____

The social lives of linguistic legacy materials

LISE M. DOBRIN & SAUL SCHWARTZ

_____

_____

EL Publishing

For more EL Publishing articles and services:

The social lives of linguistic legacy materials

Lise M. Dobrin
University of Virginia

Saul Schwartz

Abstract

Documentary linguistic data may be acquired not only firsthand, but by consulting materials that were produced by scholars, missionaries, speakers, and others in the past. Such *linguistic legacy materials* may reside in an archive or in an individual's private collection, or they may be embedded in published literature that was created for purposes other than linguistics. In this introduction to a special issue of *Language Documentation and Description*, we explore some of the reasons why linguistic legacy materials, while potentially treasure troves of evidence and insight, are nevertheless challenging to use. The main challenge, we argue, is inherent in the very nature of such materials: inasmuch as they are the products of past human meaning-making activity, they are invested with the goals, knowledge, points of view, and circumstances of those who were involved in their creation. To that extent, legacy materials can be said to possess *social lives* that originate in the past and that continue to unfold over time as they are accessed, analyzed, or put to new uses. The articles published together here tell the "biographies" of linguistic legacy materials in particular instances, drawing lessons for all who revisit and recirculate data from the past and offering perspective for documentary linguists working now to create the legacy collections of the future.

1. Introduction

Documentary linguistic data may be acquired not only firsthand, but by consulting materials that were produced by scholars, missionaries, speakers, and others in the past. Such *linguistic legacy materials* may reside in an archive or in an individual's private collection, or they may be embedded in published literature whether or not it was created for purposes of linguistics. When studied alongside more recent data, linguistic legacy materials can supply a valuable perspective on language change. Where they are the most thorough or only

extant documentation of a language that is no longer remembered or regularly used, legacy materials can be a critical resource for linguistic research and language revitalization, as well as for applied sociocultural activities like cultural renewal, land claims cases, and analysis or repatriation of artifacts.

In this introduction to a special issue of *Language Documentation and Description* on "The social lives of linguistic legacy materials", we explore some of the reasons why linguistic legacy materials, while potentially treasure troves of evidence and insight, are nevertheless challenging to use. The main challenge, we will argue, is inherent in the very nature of such materials: inasmuch as they are the products of past human meaning-making activity, they are invested with the goals, knowledge, points of view, and circumstances of those who were involved in their creation. To that extent, legacy materials can be said to possess social lives that originate in the past and that continue to unfold over time as they are accessed, analyzed, or put to new uses. "[H]uman transactions, attributions, and motivations" endow things with meanings which become "inscribed in their forms, their uses, their trajectories" (Appadurai 1986: 5). Ideals of objectivity notwithstanding, such human creations cannot be used or made sense of once extracted from their social histories without a major investment of interpretive labor – yet another difficult-to-extract, irreducibly human contribution.

The articles published together here tell the "biographies" (Kopytoff 1986) of linguistic legacy materials in particular instances, drawing lessons for all who revisit and recirculate data from the past and offering perspective for documentary linguists working now to create the legacy collections of the future. The articles trace how linguistic legacy materials acquire new meanings and transform in value as they are encountered by previously unimagined users, manipulated in unforeseen ways, and queried for new purposes. This introductory article is organized as follows. In Section 2, we show how interpretive challenges are inherent in the nature of documentary linguistic corpora despite the assumption that, if the right steps are taken, they can be made self-explanatory for future users. In Section 3, we dwell upon the notion of objectivity that undergirds this assumption, arguing that it constitutes a valued but elusive ideal in linguistics. A commitment to objectivity has become increasingly difficult to maintain in recent years as linguistic fieldwork has become more common and as linguists' human investments in their data production practices have been foregrounded in discussions of research ethics and community collaboration. In Section 4, we offer an extended case study that illustrates the limits of objectivity: the controversy surrounding the reliability of linguist Daniel Everett's corpus of Pirahã. As the documentation of endangered languages proceeds, with individual scholars producing what will often be the only significant records of a language that is no longer fluently spoken, interpretive questions like those that troubled subsequent users of Everett's Pirahã corpus are bound to be not the exception, but the norm. So it is useful to study how researchers have actually gone about interpreting records

that have become dissociated from the contextual knowledge possessed by their original authors, collectors, and audiences, as we do here. We conclude in Section 5 by encouraging documentary linguists to adopt an expressly interpretivist, rather than objectivist, orientation toward the products of their work. Finally, in Section 6 we summarize each article in the collection.

2. Linguistic legacy materials and their interpretive challenges

Contemporary researchers face nontrivial obstacles in making productive "secondary use of archival data, in which the analysis postdates the archiving" (Whalen & McDonough 2018: 53). Austin (2017) offers a thorough review of the challenges associated with legacy texts; Bowern (2003) thoughtfully discusses a number of problems she encountered using legacy texts from the 1930s in a community-based oral history project. Annotations may be in an unfamiliar form, as when texts are transcribed in an idiosyncratic orthography, or they may be otherwise inadequate for a researcher's goals, such as when the transcriptions of speech sounds are found to be inconsistent (a classic problem addressed by Boas 1889) or too coarse-grained for purposes of phonetic analysis. Legacy materials may present problems of substance (e.g., gaps, misanalyses, wrong glosses) or problems of form (e.g., incorrect or incomplete metadata, recordings that are disassociated from their transcripts). Glossing may be missing or in a language unknown to prospective users. The original handwriting may be difficult to decipher. Audiovisual recordings may require conversion out of analog or older digital formats for the materials even to be accessed (e.g., Boynton et al. 2010: 140; Czaykowska-Higgins et al. 2014). Those that are still readable may exist in outdated formats that do not conform to best practices for portability (Bird & Simons 2003; Schmidt & Bennöhr 2008).

The problems presented by linguistic legacy materials are not only technical ones. The "socio-cultural and historical context of the documents and their creation" can raise ethically troubling questions regarding intellectual property, rights, and access that call for renegotiation in the present (Austin 2017: 25). Content may be mystifying or offensive for contemporary readers (see Schwartz, this volume). Standard metadata that answers the basic who, what, where, when, and why questions about a recording may not provide the kinds of details required to guide future users in knowing how legacy materials can be responsibly accessed or circulated. Recognizing this, some authors have argued that researchers should be supplying richer metadata, including better "meta-documentation" (Austin 2013), information on the circumstances surrounding the research, and the "ideological milieu" (Innes 2010: 201), or ethnographic background, for any materials they plan to archive (see also Nathan & Austin 2004; O'Meara & Good 2010). When cast in this light, as a "far from sufficient record", legacy materials serve as a negative example that can "help us avoid making the same mistake in the record we leave behind"

(Buszard-Welcher 2010: 68). As Thieberger & Jacobson (2010: 152) observe, when "[w]orking with legacy material, we sometimes see what small additional steps researchers could have taken to make their recordings more useful". So an encounter with predecessors' data can inform our own data collection and storage practices, given that we know the materials we produce today will someday become legacy materials for our successors.

It is perhaps not much of a surprise that members of "indigenous groups […] interested in accessing materials on their languages that have been placed in language archives […] have a hard time navigating them" (Wasson et al. 2018: 239; see also Thieberger 1995). Archives are typically structured according to Western epistemological principles and embed annotations that are obscure to those unfamiliar with linguistic terminology. Communities that have tried to gather archival information on their languages have long been aware of these challenges, which stem from practices that privilege the interests of certain audiences over others, typically those with document-based rather than face-to-face knowledge systems (Barwick et al. 2019; Gibson et al. 2019). Many indigenous language activists have found that graduate training in linguistics is a prerequisite to being able to use archival resources effectively (e.g., baird 2013; Baldwin & Costa 2018). Initiatives like the Breath of Life workshops in the United States, where Native participants are mentored in the use of archival resources to support language reclamation (Baldwin et al. 2018), and projects to digitize, repatriate, and establish culturally-sensitive protocols for legacy documentation (Nathan 2013; Holton 2014; Powell 2015) reflect scholars' efforts to help mitigate the obstacles that extra-academic stakeholders face in reclaiming language documentation from archives.

But those steeped in the scholarly tradition are also stymied when they attempt to make use of archival language data. As Bowern (2018: 207) observes,

> More primary data are available, thanks to online archives. However, it needs to be acknowledged that primary data are still difficult to use without familiarity with the language (or one closely related) […]. And because there are so many languages and so few linguists working on them, it's often the case that the only person with the requisite knowledge to use raw data from a corpus collection is the linguist who collected it in the first place. So, linguists end up using the secondary sources anyway, even if the raw data are available.

Indeed, research into this problem by Wasson et al. (2018: 239) reveals that linguists find archived data "cumbersome and often frustrating to use" even though producing such data is the whole point of the language documentation enterprise. It turns out that "*none* of the intended user groups" is "able to

productively use language archives developed by linguists" (Wasson et al. 2018: 240).[1]

   As we will argue, the challenges associated with using linguistic legacy materials derive not just from unfortunate errors of omission or commission on the part of those who produced them, nor from insufficient planning for academic vs. non-academic users. In line with the position of several other authors writing about the nature of documentary linguistic materials (e.g., Lehmann 2001; Nathan & Austin 2004; Evans & Sasse 2007; Woodbury 2007, 2014; Seidel 2016), we argue that *interpretive challenges are inherent in the nature of all records produced by others at another time for their own purposes*. It is tempting to believe that future users of the documentation we produce will be spared the interpretive effort demanded of our contemporaries who are trying to decipher century-old handwritten records that employ idiosyncratic or archaic orthographies, abbreviations, or analytical terms, or those who are comparing data from multiple sources or related languages in order to reconstruct a most likely historical form. After all, the current language documentation paradigm is explicitly designed to produce records that are future-proof: transparent and accessible to the broadest possible audience of potential future users, portable, enduring, standardized, uncontaminated by our theoretical biases, and enriched with ample metadata. Yet however conscientious linguists may now be in their archiving practices, the data they produce will still present challenges for future users.

   In order to understand why, we need to return to the basic tenets of current documentary linguistics, which was founded upon the idea that data can be made self-explanatory (Himmelmann 1998). Distinguishing *documentation* – collecting, transcribing, and translating linguistic data – on the one hand, from *analysis* on the other, implies that collecting, transcribing, and translating data are not themselves analytical activities. This is what makes it conceivable that linguists could produce "a comprehensive corpus of primary data which leaves nothing to be desired by later generations" (Himmelmann 2006: 3). This "record for generations and user groups whose identity is still unknown and who may want to explore questions not yet raised at the time when the language documentation was compiled" is contrasted with records created "for a specific purpose or interest group", which are assumed to be less liable to shed light on new problems (Himmelmann 2006: 2). In this view, researchers can produce documentation that will be accessible to undefined future users by relinquishing any agenda of their own, apart from the agenda to produce a record that will

---

[1] Peter Austin (personal communication, March 2020) reports hearing from a linguist currently working with a community in the Middle East that although he had access to archival audio recordings made by another researcher working with the same group, it was more efficient to set those materials aside and carry out his own new fieldwork.

allow other users to draw on the data for their own projects. In this way, it is proposed, linguists can create transparent records that will underwrite reproducible analyses. *Transparency* is the idea that an apparatus or metadata will make records "interpretable by its future perceivers", even "useable by the philologist 500 years from now" (Woodbury 2003: 47, 2011: 160). *Reproducibility* is the idea that readers of descriptive analyses should be able to access the data on which they are based in order to "verify the results [of analysis] for themselves" (Gawne & Berez-Kroeker 2018). The emphasis in documentary linguistics on the paired ideas of transparency and reproducibility reflects the value placed on linguistic records being relatively autonomous, so that those who did not collect them will be able to use them with nearly as much expertise and contextual knowledge as those who did. Given this framework, documentary linguistics becomes a project of "feeding" self-explanatory primary data into open archives (Himmelmann 2006: 6).

Yet are recording, transcribing, and translating not themselves analytical activities (Himmelmann 1998: 162–163; see also Lehmann 2001: 88–90; Austin & Grenoble 2007: 22; Evans & Sasse 2007; Evans 2008: 346–348; McDonnell 2018)? An effort to clarify this question led the Linguistic Society of America to adopt a "Resolution recognizing the scholarly merit of language documentation" declaring the products of language documentation to be "intellectual achievements which require sophisticated analytical skills, deep theoretical knowledge, and broad linguistic expertise" (LSA 2010). In line with the LSA resolution, we argue here that compiling a documentary linguistic corpus involves analytical choices at every step of the process outlined by Himmelmann (1998: 171):

- decisions about which data to collect/include in the documentation;
- the actual recording of the data;
- transcription, translation, and commentary;
- presentation for public consumption/publicly accessible storage (archiving).

Let's consider these each in turn.

"Decisions about which data to collect/include in the documentation". Decisions about what to record are shaped by the researcher's analyses (explicit or otherwise) of the categories of communicative events within a speech community (Himmelmann 2006: 8). What is included or excluded will also reflect the agency of consultants and communities. For example, a researcher can request that speakers enact a particular kind of communicative event, but once the recording begins, speakers may put other kinds of discourse on record. Conversations cannot be controlled by the researcher or even planned by the participants; they unfold according to their own driving logic over interactional time (Enfield 2017). Even with monologic speech directed to a recording device, the communicative import and audience is often far from straight-

forward: Is the speaker communicating with the researcher? With the social group the researcher is presumed to represent? With current, past, or future members of their own community? What messages are being sent? Recovering the communicative import of a recorded narrative is a major interpretive task, not only for work with legacy materials (Silverstein 1996; Nevins 2013, 2015; Silverstein 2017; Schwartz 2019), but also when researchers revisit materials that they themselves collected and find that their consultants were responding to them in ways that they did not fully appreciate at the time of recording (Moore 2009; Dobrin 2012). Given that retrospective interpretation is often required to make sense of data that we ourselves collect suggests that interpretive challenges are inherent in nature of all records. These challenges are only amplified when working with materials produced by others.

"The actual recording of the data". When reading what someone wrote in their fieldnotes, we still do not know exactly what circumstances led the author to write down what they did. Audio recordings do not capture images. Photographs do not capture movement or sound. Even video does not capture what is happening just beyond the recording's spatial and temporal frame. When will the recording device be started and stopped? How are off-screen events influencing what is happening within the frame of the recording itself? As anthropologist Michel-Rolph Trouillot (1995: 49) emphasizes in a thoughtful meditation on the unrepresentativeness of all historical records,

> any single event enters history with some of its constituting parts missing. Something is always left out while something else is recorded. There is no perfect closure of any event, however one chooses to define the boundaries of that event. Thus whatever becomes fact does so with its own inborn absences, specific to its production.

Linguistic data is no less selective, reflecting as it does consultants' ideas about what is important or appropriate to document as well as fieldworkers' decisions and recording strategies. Researchers' goals, questions, and analyses may be relatively "facilitative", that is, they may facilitate the production of data that can readily be repurposed to answer new questions (McDonnell 2018). But even the most comprehensive language documentation will still have gaps and silences that subsequent users will need to fill in and make sense of.

"Transcription, translation, and commentary". These are all activities in which there is considerable overlap between data collection and analysis (Himmelmann 1998, 2012). The overlap reveals something of a paradox within the documentary enterprise. The most basic kind of data that fieldworkers collect are recordings of language use. But unannotated recordings "are rarely used directly as the basis for further research" because the information they contain is "too much and too complex" (Himmelmann 2012: 193). To "make the data compiled in a documentation accessible", transcriptions and translations are necessary (Himmelmann 2012: 204). Nathan & Austin (2004: 184)

warn against "a future of wading in digital quicksand – a rapidly expanding mass of digitized sound, image and video, with no way to get a foothold". Transcripts provide probably the single most important foothold. For this reason, "it is standard practice to work with a transcript of the recording which, ideally, contains all and only the [relevant] aspects of the recorded event" (Himmelmann 2012: 193).

This recognition – that documentary recordings are essentially unusable by another linguist today (let alone by limitless future audiences) without a transcript – means that on a practical level, the distinction between data and analysis simply cannot hold. Because as any documentary linguist knows, producing even the most basic transcript is a major intellectual undertaking. It is surely no accident that in a special issue of *Language Documentation and Conservation* dedicated to "Reflections on language documentation 20 years after Himmelmann 1998", Himmelmann himself chose to contribute an article about the need for linguists to better understand the act of transcription. Himmelmann cites seminal work by Ochs (1979: 44) to show how "transcription is a selective process reflecting theoretical goals and definitions". That is, transcripts are analyses and interpretations that have been shaped by the researcher's perceptions, assumptions, and aims (Bucholtz 2000), as well as by the choices and judgments of any community members or others involved in producing the transcription (see Dobrin this volume; also Haviland 1996; Urban 1996). As Himmelmann (2018: 35) puts it, "it would be rather naïve to consider transcription exclusively, or even primarily, a process of mechanically converting a dynamic acoustic signal into a static graphic/visual one. Transcription involves interpretation and hence considerable enrichment of the acoustic signal". In other words, to read a transcript is to view an analysis. It is not an unmediated encounter with raw data.

"Presentation for public consumption/publicly accessible storage (archiving)". Historical trends and archival policies shape the archived product and its metadata. Questions of credit, confidentiality, attribution, and intellectual property determine what kinds of data are held back while others are made publicly available, and in the latter case, in what forms. A degree of "editing" is done in the work of archival assembly "to make the data accessible to the uninitiated" (Himmelmann 1998: 165). For example, the segmenting of recordings into "individual communicative events" or "sessions" embeds a layer of analysis in the structure of the material (see, e.g., Himmelmann 2006: 10). What is included within the apparatus or metadata with the goal of providing other users with the kinds of knowledge the data creators/collectors possess? What contextual knowledge is impossible to make public since it is embodied in the experience of the creators/collectors and so difficult to express?

Throughout these prescribed steps for creating the most future-oriented linguistic data possible, then, analytical and interpretive choices are being made by researchers, speakers, transcribers, archivists, and other stakeholders. However committed we may be to the idea that linguistic records can be made

transparent, autonomous and agenda-less data is a mirage. Language documentation does not result in "pristine raw data" that subsequent users will be able to make sense of simply by reading, listening, or watching; rather, it is a human creation shaped by layer upon layer of analytical choices, with interpretive effort required to discern its meaning (Seidel 2016: 38). For philologists and others 500 years in the future who are trying to use the documentary materials that are being produced today, it will be a major activity to reconstruct the intellectual, social, political, and ethical milieu that gave rise to linguistics as we currently know and practice it, and to reconstruct the methods, motivations, and social relations that gave rise to any particular archival object in which they take interest.

3. Objectivity: An elusive ideal

If archived language data is not – indeed cannot be – self-explanatory, then why have documentary linguists dedicated themselves to trying to make it so? There are different ways one could try to answer this question, but the line of reasoning we would like to pursue here has to do with a set of tacit values, an "epistemological unconscious" (Steinmetz 2005), that linguists acquire as they are socialized into their discipline. We call it "the objectivity paradigm": the belief that objectivity is essential for the production of valid scientific knowledge. This scientific ideal is "knowledge that bears no trace of the knower – knowledge unmarked by prejudice or skill, fantasy or judgment, wishing or striving. Objectivity is blind sight, seeing without inference, interpretation, or intelligence" (Daston & Galison 2007: 17). Objective knowledge is asocial knowledge, separable from those who produce or use it. The notion of objectivity has been studied extensively by philosophers and historians of science; here we focus only on those key aspects that are relevant to the discussion that follows:

- Analyses should be general, and hence replicable, as opposed to being dependent upon particularities of the research situation.
- Results should be free from the effects of the researcher's personality, emotions, theoretical predispositions, desired outcomes, etc.
- The phenomena studied must exist independently of the researcher, and not come into being only by virtue of the researcher's presence (Latour 1993).
- The value derived from apprehending the systematic relations among elements abstracted from the phenomena studied eclipses everything else: the researcher's social embedding within the research site, their encounter with and experience of the phenomena under investigation, and their personal investment (aesthetic, emotional, political, etc.) in data collection and analysis (Boas 1887).

As science historians Lorraine Daston & Peter Galison (2007) have shown, the minimization, control, and even erasure of the observer in the service of making data objective has formed part of the conceptual underpinnings of western science since the mid-19th century. Of course, as an irreducibly human endeavor, the actualities of scientific practice continually challenge the objective ideal. Hence scientists find themselves drawn to engage in what sociologist of science Bruno Latour (1993) calls the work of *purification*: hiding their own tracks; keeping the hand that holds the object outside the frame when they take the picture they then take to constitute "the work".[2] That purified image is nowhere harder to maintain than in humanistic field studies, where data emerges from open-ended interactions between people whose priorities and perspectives cannot help but affect the interaction.

In linguistics, problems surrounding objectivity remained reasonably well submerged as long as one could take it for granted that language description was the most basic kind of data upon which linguistics built. But this began to change since the endangered languages issue brought language documentation into focus, legitimating fieldwork as a form of linguistic research and not just a background preliminary to it. Language documentation is often depicted as firming up the empirical foundations of linguistics as a scientific discipline, yet from the very outset, it has called the discipline's commitment to objectivity into question. Many linguists today work with languages that are contracting because their speakers are socially, economically, and/or politically marginalized (England 1992; Ladefoged 1992; Dorian 1993). Those working with such languages frequently feel called to orient their work in relation to a host of technically "non-linguistic" issues surrounding decolonization (Leonard & De Korne 2017), indigenous rights and global ecological sustainability (Nettle & Romaine 2000), community empowerment (Czaykowska-Higgins 2009), and the ethics of data collection and representation (Dwyer 2006; Austin 2010; Innes & Debenport 2010; Macri 2010; Rice 2012; Good 2018). With so much attention being given to the social dimensions of language vitality and research praxis, and with linguistic fieldwork becoming accepted as a legitimate activity

---

[2] For example, when speaking about the trajectory of their research, scientists tend to represent it in purely intellectualist terms as following from disciplinary goals and objectives. But with further probing, they will often provide accounts that make reference to "external conditions such as the availability of funding, the organization of research, and personal problems" (Yearley 1988: 352). An example of purification in language documentation would be fieldworkers adopting methods like leaving the room while the recorder is running in order to keep their presence from affecting the data. Stenzel (2014: 296, n. 14) displays an unusual resistance to purification when she refuses to "downplay the 'fun factor' in research, particularly research of the collaborative type, so heavily dependent on relationship-building and creating an ambiance in which people feel comfortable with each other".

that is frankly and regularly discussed in professional settings, it is harder than ever to keep the hand from appearing inside the frame, with both researchers and participants increasingly visible in linguistic work as people, actively shaping the record with all their biases, ambitions, and human imperfections. In Dobrin & Schwartz (2016: 259–260) we recuperate the term "linguistic social work" to refer to this characteristic of basic linguistic research.

The threat to objectivity created by linguistic social work presents something of a dilemma for documentary linguists, who are all too aware that their fieldwork is justified by the data it provides for dominant theories in their discipline. With few exceptions, these theories descend from or respond to a core program that has, since Saussure (1959[1915]), identified *langue* as the proper subject matter of linguistics. The whole goal of that program is to make linguistics scientific, i.e., subject to just the sort of objectivist assumptions we have been discussing. Indeed, ever since its founding moments, linguistics has been intensely focused on methods that can delimit language as an object of analysis independent of the situation in which it is spoken (and studied). In an insightful analysis of how linguistics simultaneously constructs both its object of study and the discipline itself, linguistic anthropologist Asif Agha recalls Bloomfield's (1926) aim to define linguistics' key object, the "utterance", in such a way that *I'm hungry* will constitute a unitary object whether spoken by a child avoiding bedtime or by a needy stranger (Agha 2007; the significance of this example is also explored in Hymes 1974: 435 ff.). "[S]tructuralist abstraction and generality" of the sort linguists are trained in requires a "distance and distancing between analyst and object of analysis" (Collins 1992: 409; see also Easton & Stebbins 2015). Having cut our teeth in a discipline founded on objectivizing structuralist assumptions that obviate both speakers and the context of speaking in this way, and with those same assumptions continuing to be held dear if not in the recesses our own hearts then in the office down the hall, it is hard to establish what place, if any, linguists and participants as persons should be understood to occupy in the activity of research – all the way down to the moments in which data is being collected.

Perhaps the closest linguists have come to acknowledging the place of the personal in documentary research is the elaboration and refinement of models of collaboration that have become a mainstay of journals and other publications where linguistic field methods are now under regular discussion. But even as this literature foregrounds important social complexities that bear on the practice of research, it still often accommodates the objectivity paradigm by turning diverse experiences or "cases" into systematic, generalizable "models" in a way similar to that by which linguists reduce speech to formal patterns and rules.

Tonya Stebbins's 2012 article "On being a linguist and doing linguistics: Negotiating ideology through performativity" presents a fascinating illustration. In it, Stebbins seeks to reconcile the tensions she faces while trying to meet the expectations of her scholarly community on the one hand, and the indigenous communities with whom she does fieldwork on the other.

Experiencing the tension as a sort of double bind, Stebbins models these sets of expectations as separate bounded spheres, each associated with corresponding interests, roles, and activities. Despairing of balancing the two spheres equally, she analyzes out a distinct "third space" to make visible the mediation between them (Stebbins 2012: 308; see Figure 1).
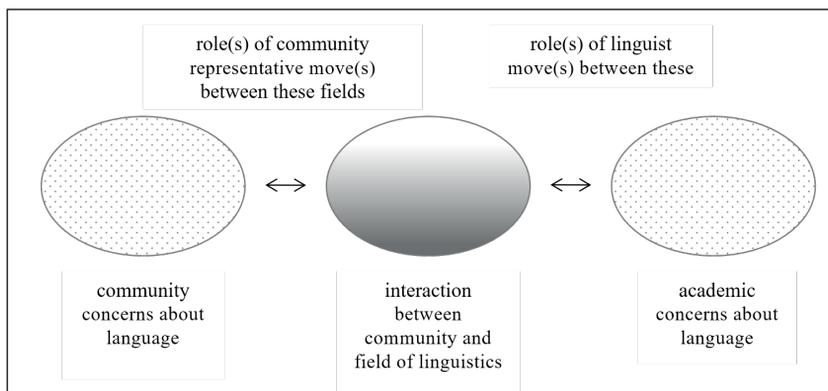

Figure 1. Stebbins's model of social relations in linguistic fieldwork

Doing this helps Stebbins (2012: 309) feel more comfortable with her personal identity in fieldwork and allows her to let go of "some of the burden of over-responsibility and guilt" she feels "in connection to 'being the linguist'". While continuing to bring her professional knowledge to her field encounters, she can disavow responsibility for community decisions, e.g., what orthography is adopted, while still acknowledging that there is some basis for her influence. The mediated model also puts community concerns on equal footing with academic ones even though the former find no obvious place within the objectivity paradigm.

Stebbins does this kind of diagramming to achieve some degree of separation between her field relationships and the substance of her research because she feels called upon by her discipline to perform an objective stance. Yet her graphical solution cannot resolve the more general problem that *field research by its very nature integrates scholarly knowledge production with personal experience*. Stebbins (2012: 293–295) adopts an "explicitly personal tone" yet worries that her experience may not "resonate […] with every reader" even as she tries to articulate a generalized model that "can be used in negotiating any kind of fieldwork arrangement". Conflicts such as this are as widespread as the objectivity paradigm itself. As David Harrison (2005: 38) notes, field linguists often struggle to justify their research in terms of a disciplinary hierarchy that values researcher objectivity and control over the contingency of fieldwork: "[W]e often feel we must apologetically present our

field data as 'anecdotal', not statistically valid or replicable, and therefore scientifically suspect or inferior".

What we see in Stebbins's article and comments such as Harrison's are scholars trying to reconcile their personal and political investments in field research with their discipline's commitment to objectivity. Interestingly, one benefit of formal collaborative protocols is that they allow linguists to insert themselves in the research process in a way that accommodates objectivist assumptions: everyone's involvement is explicitly managed through association with specific project roles. As Stebbins (2012: 309) puts it, the linguist's "influence in the community is both mediated and limited". And hence the appeal of diagrams (see also, e.g., Benedicto et al. 2007: 30; Dwyer 2010: 195; Leonard & Haynes 2010: 288) that reformulate what are ultimately very personal experiences as analytical generalizations with enduring reality across research settings and interactions.

It is the nature of linguistic fieldwork to create knowledge through direct social relations that researchers establish and nurture within speech communities, and the quality of those relations necessarily depends on factors like personality and local cultural norms that are difficult to control, let alone replicate across settings. Downplaying this while focusing on institutionalized offices or roles ("linguist", "community representative", etc.) displays just the sort of objectifying structuralist predilection that researchers have been socialized into in their study of language. In effect, the objectivity paradigm, which calls for "the countering of subjectivity" through "the suppression of […] the self", influences how linguists have conceptualized the inherently intersubjective activity of fieldwork (Daston & Galison 2007: 36). As we discuss in the next section, it is but a short step from there to our main concern, the notion that the materials resulting from linguistic fieldwork can be treated as objects that can speak for themselves.

4. The problem of representational decisions in Daniel Everett's corpus of Pirahã

Problems arising from linguists' commitment to objectivity are evident not only in the way linguists imagine the social process of fieldwork, but also in conflicts over the way linguistic knowledge is understood to be embedded in data derived from fieldwork. As Lehmann (2001: 87–88) observes,

> The documentation of [language use] requires the addition of inter-
> pretive information which is not inherent in the primary data, but
> generated by the author of the documentation. One may regret this
> necessity in the interest of objectivity. However […] objectivity in
> documentation is an illusion […]. There is necessarily a process of
> selection, which in itself is not objective and which, in fact, can be
> highly tendentious.

In this section, we look at the debate prompted by Daniel Everett's work on Pirahã as an exemplary case that brings to light tensions between the objectivity paradigm on the one hand, and the embodied nature of linguistic knowledge on the other. We show how aspects of Everett's work that have drawn criticism for having been presented in bad faith follow from the researcher's involvement as a person in fieldwork such that the resulting materials cannot unambiguously speak for themselves.

In 2005, Everett (2005: 622) published an article in *Current Anthropology* in which he argued that a "cultural constraint" restricting communication to "the immediate experience of the interlocutors" accounted for a number of remarkable features of Pirahã grammar including, most controversially, the absence of recursion. Everett framed his findings as a challenge to universal grammar, and a debate between Everett and his critics unfolded in scholarly venues like *Language* and in newspapers, magazines, blogs, and other media.[3]

Pirahã is spoken by only a few hundred monolingual hunter-gatherers living along a relatively remote tributary of the Amazon in northwestern Brazil. The only non-Pirahã people who know the language are four current and former SIL missionaries, Everett being one of them. As a result, those wanting to evaluate Everett's analysis were almost completely dependent on Everett's own account, unless they were going to spend months or years living with the Pirahã to learn and study the language themselves. Because of the obstacles to undertaking independent fieldwork with Pirahã speakers, Everett's critics were forced to rely on his data.

The controversy over Everett's analysis of Pirahã revolves around what Bucholtz (2000) calls "representational decisions", decisions about how language should be presented in written form. In this case, because the questions are about morphosyntactic interpretation, it is not just transcriptions but decisions about glossing and translation that are at issue. In an article on the conceptual relation between documentation and description, Himmelmann (2012: 194) says that representational decisions should not, in principle, be problematic:

> Segmentation and translation involve a certain amount of interpretation because neither is fully determined by the evidence available in the recording. As a consequence, two teams of research-ers working on the same recording will not produce one hundred

---

[3] Controversies like this are valuable for social studies of science because they illuminate aspects of disciplinary culture that are otherwise implicit. We discuss the Pirahã controversy here because we are interested in the assumptions about objectivity that it reveals. Although we sympathize with the reasons Everett gives for his claims, we are not attempting to determine who is right or what is true about Pirahã grammar.

percent identical transcripts/translations (though, one would hope, that the two transcripts with translation would be reasonably similar and that the differences [for example, in representing clitic items] are irrelevant for many research purposes).

The Pirahã case shows how misplaced that hope can be. The question we will be focusing on is whether Pirahã has quantifiers, specifically whether it has a word for 'all'.

Everett (2005: 624) claims that in Pirahã "there are no quantifier terms like 'all', 'each', 'every', 'most', and 'few'". In (1) we reproduce the example he provides of how Pirahã speakers express propositions that in other languages would be expressed using quantifiers:

(1)

ti	*'ogi-'aága-ó*		*'ítii'isi*	*'ogi-ó*	*'i*	*kohoai-baaí*
I	big-be(permanence)-direction		fish	big-direction	she	eat-intensive

koga	*hói*	*hi*	*hi-i*	*kohoi-hiaba*
nevertheless	small.amount	fish	intensive-be	eat-not

'We ate most of the fish.'

(lit.) 'My bigness ate at a bigness of fish, nevertheless there was a smallness we did not eat.' (Everett 2005: 624)

The form *'ogi-ó* (henceforth *'ogió*) is glossed as 'big-direction', represented in the literal translation by 'bigness'. Everett's critics, Nevins et al. (2007), responded by referencing a glossary of Pirahã quantifiers from Everett's (1983: 362) dissertation, where *'ogió* (or *xogió* in the orthography he was using at the time) is translated as Portuguese *todo* 'all' or *inteiro* 'whole, complete'.[4] Based on the glossary, Nevins et al. (2007: 37) offer their own table of Pirahã quantifiers, where *'ogió* is translated as 'all'. They also point out that in earlier work Everett (1986: 290) glosses *'ogió* in English as 'all'.

With this evidence in hand, Nevins et al. (2007: 38) return to the above example from Everett's *Current Anthropology* article and propose that his translation should be 'We were in the process of eating all the fish, but we didn't eat a (very very) small amount', replacing the translation of *'ogió* as 'bigness' with the "more straightforward" translation 'all'. In her comment on the *Current Anthropology* article, Anna Wierzbicka (2005: 641) calls Everett's glossing of this form "ludicrous" and "exoticizing". Nevins et al. (2007: 37)

---

[4] Nevins et al. cite a published version of Everett's dissertation as Everett 1987. Our references here are to Everett 1983, a copy of Everett's original dissertation that is available online.

follow suit, arguing that Everett's claim that Pirahã lacks quantifiers is "reinforced by interlinear glosses that, [unlike those of his earlier work], resolutely reflect the non-quantificational uses of quantificational morphemes, and 'literal' translations that […] yield gibberish". In other words, they charge Everett with manipulating his glosses to justify his interpretation.

Everett's critics consider his earlier work more reliable, in part because it "presents Pirahã examples in the standard manner familiar in linguistics, with dashes and spaces separating morphemes judged to be independent, and with glosses and translations offering a fair approximation of the meaning of Pirahã rendered into Portuguese […] and English" (Nevins et al. 2007: 7). In his later work, by contrast, they find Everett's "morpheme division […] more extreme" with his glosses "appear[ing] to reflect etymology rather than current meaning"; for example, where earlier work glossed a Pirahã word as 'hammock', later work glossed it as 'cloth arm' (Nevins et al. 2007: 7).

To these criticisms Everett (2007: 20) responded by saying, "Yes, folks, I *did* think that those were quantifiers 25 years ago. But I do not now". In fact, he says, "I have learned a lot about the morphology that I did not know" at the time of the earlier writings, in part owing to the fact that he "spent an additional fifty-three months in the field since that time" (Everett 2007: 7). Whereas Everett's critics consider his earlier work to be more reliable than his later work, Everett not unreasonably frames his own trajectory as moving toward greater understanding. Much of his early work, he explains, employed morpheme breaks and glosses that he "inherited from [his missionary] predecessors among the Pirahãs", but "in the intervening years I have seen that almost all of these morphemes are in fact strings of smaller morphemes and that my original analysis was too coarse-grained" (Everett 2009: 424). Everett presents his later glosses as more accurate and his earlier glosses as due to his then more limited knowledge. He says, "I would be failing to express the richness and distinctness of Pirahã if I were to overtranslate just so that the free translation could look less strange" (Everett 2007: 20).

For simplicity we have focused on an example related to the semantics of quantifiers rather than on recursion, the topic for which this debate is best known, but the same dynamics characterize the debate on both.[5] We are also

---

[5] Debate surrounding the suffix *-sai*, for example, displays the same dynamics as those seen with quantifiers that we discuss here. Everett (2005: 629 and earlier work) glosses *-sai* as a nominalizer, on which basis Nevins et al. (2007: 15) suggest that Pirahã is one of many languages that use constructions with nominal properties to mark subordinate clauses. In response to his critics, Everett (2007: 10) claimed that in the time since his earlier work he "discovered that *-sai* in fact marks old information and is not a nominalizer at all". From 2007 onwards, Everett (2007: 10–14, 2009: 408–412, 2010: 9–13) glosses *-sai* as 'old information', a "reanalysis" (Everett's term) that his critics characterize more pejoratively as a "retraction" (Nevins et al. 2009: 673). As with the debate about quantifiers, the debate about *-sai* hinges on when and how data can be

looking at data presented in published sources rather than documentary materials as such, but the problem they present is the same. The question these controversies raise is an important one for a theory of documentation that has as its "central task […] making description accountable and replicable" (Himmelmann 2012: 199): Can other researchers who doubt Everett's claims use his representations of Pirahã data to evaluate them? Everett's answer is no, because they lack his intuitions as a speaker of the language acquired through his long experience in the field. He says, "I am bemused as to why people who have never learned a single word of a language feel at liberty to suggest translations for it, based on nothing more than their expectations from other languages" (Everett 2007: 13; see also Everett 2007: 18, 26, n. 14).

Everett (2007: 12, 30) contrasts his fieldwork-based proficiency in Pirahã with his critics' "eyeballing", which he derides as a form of "pseudo-research" that has led them to make "embarrassing speculations". Interestingly, these kinds of statements appear repeatedly in a manuscript Everett posted in 2007 to the open access repository LingBuzz but are absent from the official version he ultimately published in *Language*. Somewhere in the transition from gray literature to flagship journal, the appeal to authority based on the linguist's own fieldwork experience and resulting speaking ability fell away (see also Everett 2010: 1, n. 1). On the one hand, it is hard to discount the embodied knowledge of a linguist who spent a total of *seventy-four months* in monolingual immersion fieldwork. Would such a person not be in a privileged position to gloss and translate sentences in the language? On the other hand, this kind of direct appeal to the linguist's intuition fits uneasily with the idea that the data should stand on its own and be clearly dissociable from the person who collected it, as the objectivity paradigm demands. As Chelliah & de Reuse (2011: 376) state flatly in their comprehensive *Handbook of linguistic fieldwork*, "the fieldworker cannot rely on his/her own judgments even if s/he becomes as fluent as a near-native speaker". And some authors, such as Coulmas (1981: 9 ff.) writing in his introduction to *A festschrift for native speaker*, question whether it is justified to use the self as a source of data even if one *is* a native speaker.[6]

We should clarify the assumptions underlying the two positions. Everett's claim, based on his field experience and language proficiency, is that he is specially positioned to interpret his data. In general, Everett (2004: 144–145) argues, "there is really no one better placed to interpret data from a theoretical perspective than the fieldworker who collected" it, since that person can draw

---

interpreted by others, given that fieldworkers' representational decisions may change over time.

[6] Our informal observations lead us to believe that this view is not uncommon among typologically-oriented linguists.

on their "own history and person" to inform their "observations and conclusions". In other words, according to Everett (2004: 144), interpreting data depends on experience, and "since no two experiences will be exactly the same, no object can be the same to two people (or for one person at two times). This includes grammars and the other outputs of fieldwork". From this perspective, Everett's early documentation is a different object to him today than it was to him when he produced it, just as it is a different object to Nevins et al. than it is to Everett, past or present. The other position, which reflects the objectivity paradigm, asserts that fieldworkers have no special authority to interpret (or reinterpret) their materials because data is dissociable from the researcher who collected it. Once published, data has entered the scholarly record, and subsequent users can safely treat it as a stable autonomous object. From this perspective, what matters is not experience but "robustness of evidence and soundness of argumentation", that is, "whether conclusions follow from their premises, whether published data has been properly taken into account and whether relevant previous research has been represented and evaluated consistently and accurately" (Nevins et al. 2009: 680).

Of course, the idea that documentary recordings and fieldnotes should have integrity as objects that stand on their own is what motivates the intense investment linguists are making at present in endangered language archiving. In order to maximize coverage given the urgency of the endangerment problem, students are encouraged to "choose an undescribed language" as the target of documentation, or at least to "study the same language in another location, with different native speakers, who […] speak another dialect" (Chelliah & de Reuse 2011: 82, 84). The disciplinary future we are setting ourselves up for will thus put us squarely in the same position we are in when trying to evaluate Everett's Pirahã data: What is attributable to the language, and what is attributable to the researcher's personal investment in the glossing – not to mention the selection of speakers, the precise sample of speech that was captured, the moment at which the recorder turns on and off, and so on?

Analysis of documentary material is something that evolves. Few field-workers would assume that their first (or even second) analysis was necessarily the best. They know their representations are unstable: transcriptions, morpheme breaking, and glossing change over time. Evans & Sasse (2007: 60) liken the process of annotation to an "endless quest" that "gradually leads to a better understanding of the utterances under study". Given that annotations "may continue to be worked through for [an] in principle unbounded time", their instability must be considered a systematic feature of language documentation, rather than a bug (Evans & Sasse 2007: 73; see also Jung & Himmelmann 2011: 204). Like Everett, Evans notes that he revises his glosses as he learns more about the language. For example, early texts might have a prefix glossed 'Realis' that later texts gloss as 'Assertive' (Evans & Sasse 2007: 76).

As the Pirahã case shows, such seemingly minor changes may present considerable difficulties for subsequent users even though in our own materials

we tend to read *through* them. This is because we read our own materials in conjunction with context that we know from personal experience and that informs our interpretation of the data. The literature on fieldnotes in anthropology refers to such contextual information as "headnotes" (Sanjek 1990), knowledge that is stored in fieldworkers' memories. The standard metadata categories cannot capture much of the information in headnotes, which is nevertheless triggered whenever researchers revisit their own data. As Thieberger & Berez (2012: 117) phrase it, "there is more to research on a previously undescribed language than just the recorded data. The headnotes, the associated memories, and contextualization of the research that remain in the head of the researcher continue to inform their analysis". In the case of elicitation, for example, we may remember how the speakers were positioned, who was present in the background, what we were attempting to get at, how expected or surprising a response was, idiosyncrasies of the speaker, etc. McDonnell (forthcoming; see also Bishop 2006) proposes that linguists enrich their transcripts with "pragmatic annotation" – contextual and ethnographic commentary – in an attempt to capture what is in the researcher's headnotes. While taking this step certainly adds to the richness of the material, the reality is that context is endless. As with metadata, pragmatic annotation can only capture a small fraction of what a researcher prospectively imagines an unspecific future user would want to know.

Everett's early work provides a case in point. Presenting an apparently grammatical Pirahã noun phrase with multiple recursively embedded modifiers, *kabogáohoi biísi hoíhio xitaíxi* 'two heavy red barrels', Everett (1983: 132, 1986: 273) noted that the example was "rather artificial" since it was "not taken from textual material but rather was separately elicited". More than twenty years later, Everett revisited this example in the context of the debate about recursion in Pirahã and explained that he himself had originated the sentence. Everett had set up two heavy red barrels and said, "*kabogáohoi biísi hoíhio xitaíxi*", asking Pirahã people if they could say that about them. A number of consultants said, "Yes, you can say that", though he could only get a few of them to repeat the phrase themselves, and most would not (Everett 2009: 422). Years later, when Everett asked these same consultants why they did not say things like *kabogáohoi biísi hoíhio xitaíxi*, they said, "Pirahãs do not say that". When Everett (2009: 422) reminded them about the scenario with the red barrels from years earlier he was told, "You can say that. You are not Pirahã".[7]

---

[7] Everett provided an oral account of this example during a 2006 talk at MIT (Everett 2006 at 42:10–55). The larger lesson that Everett derives from this incident, that the examples in grammars cannot be taken at face value, is in accord with our assertion that fieldworkers possess knowledge they use to make sense of their own data that others will not have.

Needless to say, none of this context would ever have come to light if the question of recursion in Pirahã had not become theoretically pressing and prompted him to revisit it. Nor is it clear how subsequent users of Everett's (or any linguist's) data would be able to reconstruct such information without being able to talk to the fieldworker and/or their consultants. It is the absence of such "back-story" headnotes, embodied in people's memories but not directly accessible from an archive, that cause interpretive challenges for researchers working with legacy materials.[8]

In contrast to linguistics, where adherence to the objectivity paradigm makes it seem self-evident that previously collected data could be unproblematically used by linguists and others in the future, anthropologists generally recognize that their recordings and fieldnotes are incomplete without the researcher's embodied knowledge. The American Anthropological Association's "Statement on the confidentiality of field notes" suggests that to treat them as transparent data "is to view them in an incorrect light [that] distorts their true nature and utility" (AAA 2003). Anthropologist Lisa Cliggett (2016: 236) points out that for the researchers who produced them, "digging through forty years of […] fieldnotes may seem like visiting an old friend. [But f]or a newcomer […] it can feel like an impossible task". Indeed, she says, it "can be more time-consuming and intellectually challenging than doing one's own field research" (Cliggett 2016: 236).[9] A disciplinarily honed suspicion of objectivity and acknowledgment of headnotes is part of the reason that archiving has never taken off in anthropology like it has in linguistics; as the Pirahã controversy shows, it is not that there is an inherent difference in the nature of the materials. Although some anthropologists do advocate for donating fieldnotes to libraries and archives (Silverman & Parezo 1995), and there are objectivist-comparativist projects like the Human Relations Area Files (Tobin 1990), those are somewhat countercultural within anthropology. Instead, "anthropologists have a well-established convention of retaining their fieldnotes as a personal archive" (Lederman 2016: 259). The reason for this is that, as Rena Lederman (2016: 261) puts it,

---

[8] It is not only the products of elicitation that present challenges for later interpretation. Even with texts, which might seem to be linguistic objects with a reasonable degree of autonomy, the researcher's entanglement in their materials can be dramatic. When Dobrin (2012) decided to analyze an Arapesh text in her collection – for no more significant a reason than that it was short and she needed a handy model to use in a class – she found herself deeply implicated in just about every aspect of "the data" from the topic the speaker chose to address, to where the line breaks were in the transcription, to the flow of the narrative discourse, to where the story appeared to begin and end.

[9] As noted above, linguists say similar things.

> [n]o matter how detailed the contextual information provided is, secondary users are likely to need not so much something more as something else […]. [E]mbodied experience substantively distinguishes primary field researchers from secondary users of research records […]. [M]etadata can compensate for thin notes but is a poor proxy for living memory, which is not an archivable artifact.

After having the unusual experience of writing an ethnography of a culture using another anthropologist's fieldnotes, Nancy McDowell (1991: 25) concluded that "no anthropologist can ever take the notes of another and make of them what the original author might have" due to "what never enters the notebooks – the contexts, the connections and associations, the sensations, and even the obvious, most of which the original researcher probably could recall, probably does know, or at least would use as framework in writing". Due to their emphasis on headnotes, anthropologists generally set aside questions of replicability or reproducibility even within the tradition of cultural "re-studies" (Hammersley 2016).[10]

In insisting that querying his memories and intuitions as a Pirahã speaker is reasonable given his long experience of immersive language learning, Everett is rejecting the objectivity paradigm and staking out a place for embodied knowledge, i.e., headnotes, in research with documentary linguistic materials. The fact that his critics found this so maddening – as if Everett were breaking an unspoken ground rule about how research is supposed to work – shows how deeply ingrained the objectivity paradigm is in linguistics. But the fact that linguists today also express frustration that the archived documentation they are so arduously producing isn't getting much secondary use suggests that it may be time to rethink the assumption that objectivity is required for scholarly knowledge to be valid or of value.

The embedding of persons in linguistic data has no ready solution to the extent that headnotes are the ineluctable complement to recordings and fieldnotes. But headnotes are to some degree sharable through interaction, so that working with someone else on their materials can productively serve as an intermediate experience between working on one's own materials, where the headnotes and embodied knowledge are always there, and interpreting someone else's archive, where the data (plus metadata) have to stand alone. This is something we have taken advantage of in our own work: Dobrin in

---

[10] This is part of why linguistic anthropologists working with legacy texts have invested so much effort in recovering the interactional dimensions and communicative intentions behind them (Silverstein 1996; Nevins 2013, 2015; Silverstein 2017), as well as in analyzing the conditions of text production more generally (Voegelin 1952; Darnell 1990; Stocking 1992: 60–91; Bauman and Briggs 2003: 255–298; Silverstein 2015; Epps et al. 2017).

collaboratively glossing Robert Conrad's Bukiyip texts to prepare them for archiving (see Dobrin, this volume), and Schwartz in interviewing Jimm Goodtracks about his 1970s recordings of Chiwere-speaking elders, recordings that Goodtracks (relying on his headnotes) interprets as reflecting the late elders' "traditional" attitudes toward bodies and sexuality (see Schwartz, this volume). It is also something the Linguistic Society of America's Committee on Endangered Languages and their Preservation has tried to support by pairing senior researchers with students interested in working with them to archive their documentary collections. Collaborative study of previously collected material undertaken jointly by collectors/creators and current (not just anticipated, imagined, or potential) users embeds the materials in a social process that begins to identify and bridge the gap between what the creator/collector knows *but would never think to say*, and what aspects of the material other users actually wish to better understand. Such dialogue makes a virtue out of what, from the perspective of the objectivity paradigm, might seem like a defect: that the products of research come with people and past interactions embedded in them. Our experiences working with original collectors to better understand their materials have taught us that it is important to approach the task with curiosity, rather than in an evaluative mode in which the goal is to assess whether or how the original collector might have "gotten it wrong". Joint study of an original collector's recordings, notebooks, preliminary analyses, and works-in-progress presents what is all too often a singular opportunity for another person interested in the materials to probe and explore not only *what* the materials are, but also how the original collector's ways of thinking and past actions help explain *why* the materials take the forms that they do.

Even when it is impossible to consult with the original collector of legacy materials, it may still be possible for researchers to come to a richer understanding of others' documentation by engaging with consultants, source communities or their descendants, and historical records that may shed light on the interactional and cultural contexts that shaped the data and gave it significance across the trajectory of its social life. In other words, what linguists need is, as Lederman puts it, "not just something more, but something else" – not just more metadata or better archiving practices, but an alternative theorization of how to make use of the knowledge that archives, together with those involved in their creation, jointly "contain".

5. Conclusion: An interpretive approach to documentary linguistic materials

The alternative theorization that we have been working toward in this article calls for documentary linguists to let go of the objectivity paradigm and align themselves with *interpretivist* researchers in the social sciences who share an "appreciation for the centrality of meaning in human life" and who are

committed to "a reflexivity on scientific practices" (Yanow & Schwartz-Shea 2014: xiv; see also Geertz 1973, Rabinow & Sullivan 1987; Steinmetz 2005; Schwartz-Shea 2014; Bevir & Blakely 2018). Interpretivism takes for granted that people are always engaged in sense-making as they act in a world shaped by human agency, intersubjective understandings, and social structures. From an interpretivist perspective, "the humanity of the researcher [is not denied or set aside, but rather] actively theorized" because once we acknowledge that researchers themselves exist within the same social field as the objects they study, "there is no position outside of society from which the scientist can 'objectively' observe" (Schwartz-Shea 2014: 3).

In a classic article on English speakers' assumptions about how communication works, linguist Michael Reddy (1979: 309) queries the commonsense belief that the "books, and tapes, and films, and photographs" in libraries and repositories actually "contain" the linguistic and cultural heritage of the societies they come from. Standard objectivist models of language documentation and archiving similarly imply that "the more signals we can create and preserve, the more ideas we 'transfer' and 'store'" (Reddy 1979: 310). But as Reddy (1979: 309) points out, "there are no ideas whatsoever in any libraries. All that is stored in any of these places are odd little patterns of marks or bumps or magnetized particles capable of creating odd patterns of noise". To use any kind of historical records, sense-making humans need to reconstruct for themselves, in their own historical moment, the "patterns of thought or feeling […] which resemble those of [the] intelligent humans" whose activities the materials record (Reddy 1979: 309). "All that is preserved" in a library or archive, Reddy (1979: 309–310) reminds us, is the "opportunity to perform this reconstruction". In a world where data is expected to speak for itself, the task of reconstruction can only ever seem remedial, compensating for an imputed insufficiency. But linguistic legacy data can never be so self-explanatory that those involved in its creation can be disregarded, or the task of reconstruction is rendered unnecessary. Anyone working with legacy materials – linguists, communities involved in reclamation and revitalization efforts, researchers from other disciplines, or others – will have to interpret what those materials meant to their creators, what new meanings they might take on in the context in which they are being used, and what roles they themselves as persons might play in the materials' circulation and reception.

We hope the recognition that using others' documentary materials will always require reconstructive effort can ease some of the intense pressure that documentary linguists are under to get everything right for the sake of the future, because the task is endless, and so, hopeless. The quality of our recordings may not always be excellent. The exigencies of note-taking mean that field notebooks will sometimes be messy or in places incomplete or unclear. Although documenters might aim for representativeness, certain topics, speakers, linguistic forms, and genres will be over- or underrepresented in our collections. Constraints of time and knowledge will prevent us from

transcribing and glossing as much, as thoroughly, or as accurately as we would wish. Everyone's knowledge is only partial, and energy and attention are limited. Metadata will be incomplete and sometimes wrong, despite the very best of intentions. Producing documentary materials is human work, subject to human limitations. So the materials will call upon future users to do the similarly human work of interpreting them. Of all the clues that we can leave for those future users, information that helps explain *why* the materials are the way that they are is among the most helpful. As we hope to have shown here, that information will often be about the persons involved in the materials' creation. So for collectors/compilers to hold back or filter out traces of themselves – actions they took, how they felt, what they were thinking – in an effort to make their documentary materials more objective, or focused entirely on the topics studied, can only be counterproductive.

By exploring what happens when we make use of others' data, the articles in this collection provide insight into how future audiences will use the documentation that we are producing today. As we will see, many important aspects of context are never recorded either in the documentation or the metadata, in part because what is considered important changes over time, and in part because the exact context that matters depends on subsequent users' knowledge and interests. In attempting to fill in those gaps we are brought into new or renewed engagements with other linguists, consultants, curators, and stakeholders. Which is just what we should expect once we stop trying to insist that data be objective and accept that it is at bottom social: "To the extent that our […] questions begin with the fact that both we and our interlocutors act, think, hope, remember, foresee, and form judgements amid a world of other people, our engagements should return us to them again" (Keane 2005: 85).

6. Summary of contributions

The first two articles in the collection illustrate the kinds of efforts that can be required to reconstruct missing contextual information about documentary linguistic records – as well as the potential rewards of making those efforts. Lise Dobrin describes her unexpected inheritance of a corpus of text transcripts in Bukiyip, a variety of the Arapesh language that is closely related to the one on which she had previously conducted fieldwork in Papua New Guinea. When audio recordings corresponding to some of the Bukiyip texts later came to light, Dobrin discovered that certain features of the transcripts could not be attributed to the documented speech, as she had assumed, but were creative interventions made by the original (unknown) Bukiyip transcriber. The discrepancies between transcripts and audio turned out not to be entirely disadvantageous, however, as they provided an additional angle from which to interpret the materials. They allowed insights into the perspectives of those involved in the materials' production, and they even revealed linguistic properties of the

language that could not be discerned from either text or audio alone. A transcript that diverges from an audio recording is therefore not necessarily wrong but may reflect other criteria or knowledge the transcriber brought to bear on the task that can be recovered through comparison and interpretation.

Tobias Weber discusses his work with archival materials documenting the Kraasna dialect of South Estonian that were created by folklorists over the past century. While Weber began studying the materials out of an interest in Estonian dialectology, he found he could not use them without first reconstructing the motives and methods of those involved in the materials' production; he also had to unravel the dependencies that existed between subsets of the materials. Doing this turned out to be a major research project in its own right, but one that led to the production of new knowledge. Weber rediscovered forgotten archival recordings and encountered materials that allowed him to date the death of the last native speakers of Kraasna to the mid-1930s. He also gained insights into the work procedures and attitudes of the earlier researchers that bear on the materials' interpretation; for example, some folklorists transcribed and analyzed Kraasna speech in a way that supported their goal of a promoting a unified standard for the Estonian language. Weber's philological approach draws attention to the human factors in language documentation, including those introduced by researchers themselves.

Sam Beer explores the circumstances that resulted in the near loss to scholarship of materials documenting the Ugandan language Soo that were produced by John M. Weatherby in the 1960s. The story Beer tells is an edifying one because Weatherby tried repeatedly over his career to make his materials useful to other researchers; moreover, the preservation and sharing of oral data was an acknowledged priority in the field of African history in which he worked, just as it is in linguistics today. Beer explores the problem of Weatherby's materials' un-discoverability in light of the differing – and changing – attitudes toward texts that have been held by the disciplines of history, linguistics, and anthropology over time. Beer's study makes it clear that while projects like documenting and archiving might seem logical and obvious while they are underway, they are in fact disciplinarily and historically contingent. Beer concludes by reflecting on the importance of an additional human factor relevant to the circulation of linguistic legacy materials, namely the web of social relations in which producers and users of the materials are embedded. For Beer, formal archival channels were insufficient, and he only gained access to Weatherby's Soo materials through his personal and professional relationships.

Among the human factors that imbue linguistic legacy materials with meaning are the ideas that inform their community reception. In his contribution to the collection, Josh Wayt discusses contemporary Dakota speakers' appreciation for the unusual linguistic register of the Dakota Bible. Given the way it was composed – translated by French missionaries in collaboration with a French-Dakota bilingual fur trader in the 1800s – the

Dakota Bible is filled with calques and dependent-marked structures that are not typical in Siouan languages. But rather than interpreting this style as impure or "bad Dakota", as many linguists do, fluent speakers find it appealing and appropriate for a religious text. Why would this be? Wayt argues that Dakota speakers read the Bible's unusual linguistic qualities in light of an ethos of speaking that values active and astute listening, using one's interpretive skills to listen or read past the surface of what is said and discern its hidden meanings. Astute listening is a hallmark of communicative competence in Dakota, and indeed something sophisticated Dakota speakers regularly offer their interlocutors the opportunity to exercise through use of artful speech forms like puns and word play. Wayt thus shows that the quality of documentary materials cannot be understood apart from the interactions that gave rise to them and the cultural values of those who use them.

Like Wayt, Saul Schwartz examines how legacy materials are received and interpreted by contemporary community audiences. In Native North America, language documentation and revitalization often goes hand-in-hand with cultural reclamation efforts. While legacy materials constitute an invaluable repository of information about cultural heritage, they can also include material that is confusing or even offensive for current audiences. Schwartz describes how Chiwere and other Siouan language activists seek to manage their "cultural face", or protect the image of the community's traditional culture, when they represent sexual and scatological incidents in narratives or address the topic of "bad words". However, there is little consensus among community linguists about how to handle such material, leading to contrasting interpretations of legacy data, controversial concealment of sensitive information, and accusations of unwarranted censorship and gatekeeping. Given this lack of consensus, the problem of swearing and obscenity cannot be solved by restricting access to a specific social group within a community, as is sometimes done with sacred material. Instead, Schwartz argues, the problem must be acknowledged and addressed as an integral part of the interpersonal and cross-cultural setting in which documentary linguistic materials were produced in the past and acquire new meanings today.

Sean O'Neill and Saul Schwartz also explore community reception of legacy materials, starting from the observation that words are the most salient aspect of language to many speakers and learners. In communities where people view language shift and reclamation through a lexical lens, it shapes what they consider to be important or valuable in legacy materials as well as local goals for language documentation and revitalization. Drawing on his experiences collaborating with the Plains Apache, Hupa, and Ponca communities, O'Neill describes how words drawn from legacy materials have important social value for Native Americans today: in naming ceremonies and restoring vocabulary for semi-speakers among Plains Apaches, in recovering traditional Hupa worldviews and historical experiences, and in constituting a Ponca cultural and political identity distinct from that of the closely related Omaha. Schwartz

develops some of the implications of these examples for language revitalization. While it is often assumed that producing new fluent speakers should be the ultimate goal of language revitalization, there are numerous revitalization programs where learners do not become fluent, but they do know and use words, which gives them a sense of social and cultural connection to their community. Knowledge of words, and the stories and metalinguistic discourses associated with them, should be better appreciated as a legitimate and important outcome of language revitalization for many communities.

The articles published here thus offer a range of perspectives and possibilities for studying the social lives of linguistic legacy materials. From one angle, such materials invite questions about origins. By engaging with their creators and delving into related archival collections, it is often possible to recover and illuminate the complex social contexts that gave shape to the materials, from the disciplinary agendas that led to the research practices employed, to the communicative goals that speakers and consultants brought to the encounters recorded. From another angle, the ways that legacy materials are received and recontextualized by new audiences raise questions about meaning that can be explored by tacking back and forth between archives and work with contemporary scholars and communities. Of course, as the articles show, questions about origins may lead into questions about reception – and vice versa. Most importantly, the articles illustrate and insist that for those trying to make sense of them today, linguistic legacy materials are at once objects and products of interpretation, just as they were for the linguists, speakers, consultants, and communities who were involved in creating them.

Acknowledgments

Southeast Regional Linganth Exchange in Charlottesville, VA, in 2019. Finally, we want to acknowledge the special debt we owe to Rena Lederman. Her deep study of how knowledge gets produced in the humanistic social sciences has been an inspiration and a model for us both.

References

Agha, Asif. 2007. The object called 'language' and the subject of linguistics. *Journal of English Linguistics* 35. 217–235.

American Anthropological Association. 2003. Statement on the confidentiality of field notes. https://www.americananthro.org/ParticipateAndAdvo cate/Content.aspx?ItemNumber=12946 (accessed 2021-12-23).

Appadurai, Arjun. 1986. Introduction: Commodities and the politics of value. In Arjun Appadurai (ed.), *The social life of things: Commodities in cultural perspective*, 3–63. Cambridge: Cambridge University Press.

Austin, Peter K. 2010. Communities, ethics and rights in language documentation. *Language Documentation and Description* 7. 34–54.

Austin, Peter K. 2013. Language documentation and meta-documentation. In Mari Jones & Sarah Ogilvie (eds.), *Keeping languages alive: Documentation, pedagogy, and revitalisation*, 3–15. Cambridge: Cambridge University Press.

Austin, Peter K. 2017. Language documentation and legacy text materials. *Asian and African Languages and Linguistics* 11. 23–44.

Austin, Peter K. & Lenore A. Grenoble. 2007. Current trends in language documentation. *Language Documentation and Description* 4. 12–25.

baird, jessie little doe. 2013. How did this happen to my language? In Leanne Hinton (ed.), *Bringing our languages home: Language revitalization for families*, 19–30. Berkeley: Heyday Books.

Baldwin, Daryl & David J. Costa. 2018. Myaamiaataweenki: Revitalization of a sleeping language. In Kenneth L. Rehg & Lyle Campbell (eds.), *The Oxford handbook of endangered languages*, 553–570. Oxford: Oxford University Press.

Baldwin, Daryl, Leanne Hinton, & Gabriela Pérez-Báez. 2018. The Breath of Life workshops and institutes. In Leanne Hinton, Leena Huss, & Gerald Roche (eds.), *The Routledge handbook of language revitalization*, 188–196. New York: Routledge.

Barwick, Linda, Jennifer Green, Petronella Vaarzon-Morel, & Katya Zissermann. 2019. Conundrums and consequences: Doing digital archival returns in Australia. In Linda Barwick, Jennifer Green, & Petronella Vaarzon-Morel (eds.), *Archival returns: Central Australia and beyond*, 1–27 (Language Documentation & Conservation Special Publication 18). Honolulu: University of Hawai'i Press.

Bauman, Richard & Charles L. Briggs. 2003. *Voices of modernity: Language ideologies and the politics of inequality*. Cambridge: Cambridge University Press.

Benedicto, Elena, Demetrio Antolín, Modesta Dolores, M. Cristina Feliciano, Gloria Fendly, Tomasa Gómez, Baudilio Miguel, & Elizabeth Salomón. 2007. A model of participatory action research: The Mayangna linguists' team of Nicaragua. In Maya Khemlani David, Nicholas Ostler, & Caesar Dealwis (eds.), *Working together for endangered languages: Research challenges and social impacts*, 29–35 (Proceedings of FEL 11). Bath: Foundation for Endangered Languages. https://web.ics. purdue.edu/~ebenedi/papers/Benedicto_FEL2007.pdf (accessed 2021-12-23).

Bevir, Mark & Jason Blakely. 2018. *Interpretive social science: An anti-naturalist approach*. Oxford: Oxford University Press.

Bird, Steven & Gary Simons. 2003. Seven dimensions of portability for language documentation and description. *Language* 79(3). 557–582.

Bishop, Libby. 2006. A proposal for archiving context for secondary analysis. *Methodological Innovations Online* 1(2). 10–20.

Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Language* 2(3). 153–164.

Boas, Franz. 1887. The study of geography. *Science* 9(210). 137–141.

Boas, Franz. 1889. On alternating sounds. *American Anthropologist* 11. 47–53.

Bowern, Claire. 2003. Laves' Bardi texts. In Joe Blythe & R. McKenna Brown (eds.), *Maintaining the links: Language, identity and the land*, 137–143 (Proceedings of FEL 7). Broome: Foundation for Endangered Languages.

Bowern, Claire. 2018. Reflections on linguistic fieldwork. In Bradley McDonnell, Andrea L. Berez-Kroeker, & Gary Holton (eds.), *Reflections on language documentation 20 years after Himmelmann 1998*, 202–209 (Language Documentation & Conservation Special Publication 15). Honolulu: University of Hawai'i Press.

Boynton, Jessica, Steven Moran, Helen Aristar-Dry, & Anthony Aristar. 2010. Using the E-MELD School of Best Practices to create lasting digital documentation. In Lenore A. Grenoble & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 133–146. Amsterdam: John Benjamins.

Bucholtz, Mary. 2000. The politics of transcription. *Journal of Pragmatics* 32. 1439–1465.

Buszard-Welcher, Laura. 2010. Necessary and sufficient data collection: Lessons from Potawatomi legacy documentation. In Lenore A. Grenoble & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 67–74. Amsterdam: John Benjamins.

Chelliah, Shobhana L. & Willem J. de Reuse. 2011. *Handbook of descriptive linguistic fieldwork*. Dordrecht: Springer.

Cliggett, Lisa. 2016. Preservation, sharing, and technological challenges of longitudinal research in the digital age. In Roger Sanjek & Susan W. Tratner (eds.), *eFieldnotes: The makings of anthropology in the digital world*, 231–250. Philadelphia: University of Pennsylvania Press.

Collins, James. 1992. Our ideologies and theirs. *Pragmatics* 2. 405–415.

Coulmas, Florian. 1981. Introduction: The concept of native speaker. In Florian Coulmas (ed.), *A festschrift for native speaker*, 1–25. The Hague: Mouton.

Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian Indigenous communities. *Language Documentation & Conservation* 3(1). 15–50.

Czaykowska-Higgins, Ewa, Martin D. Holmes, & Sarah M. Kell. 2014. Using TEI for an endangered language lexical resource: The Nxaʔamxcín database-dictionary project. *Language Documentation & Conservation* 8. 1–37.

Darnell, Regna. 1990. Franz Boas, Edward Sapir, and the Americanist text tradition. *Historiographia Linguistica* 17(1–2). 129–144.

Daston, Lorraine & Peter Galison. 2007. *Objectivity*. New York: Zone Books.

de Saussure, Ferdinand. 1959[1915]. *Course in general linguistics*. Charles Bally & Albert Sechehaye (eds.). Wade Baskin (trans.). New York: McGraw-Hill.

Dobrin, Lise M. 2012. Ethnopoetic analysis as a methodological resource for endangered language linguistics: The social production of an Arapesh text. *Anthropological Linguistics* 54(1). 1–32.

Dobrin, Lise M. & Saul Schwartz. 2016. Collaboration or participant observation? Rethinking models of "linguistic social work". *Language Documentation & Conservation* 10. 253–277.

Dorian, Nancy C. 2003. A response to Ladefoged's other view of endangered languages. *Language* 69(3). 575–579.

Dwyer, Arienne. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In Jost Gippert, Nikolaus P. Himmelmann, & Ulrike Mosel (eds.), *Essentials of language documentation*, 31–66. Berlin: Mouton de Gruyter.

Dwyer, Arienne M. 2010. Models of successful collaboration. In Lenore A. Grenoble & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 193–212. Amersterdam: John Benjamins.

Easton, Catherine L. & Tonya N. Stebbins. 2015. On becoming an object of study: Legitimization in the discipline of linguistics. In Rik De Busser & Randy J. LaPolla (eds.), *Language structure and environment: Social, cultural, and natural factors*, 319–352. Amsterdam: John Benjamins.

Enfield, N. J. 2017. *How we talk: The inner workings of conversation.* New York: Basic Books.

England, Nora C. 1992. Doing Mayan linguistics in Guatemala. *Language* 68(1). 29–35.

Epps, Patience L., Anthony K. Webster, & Anthony C. Woodbury. 2017. A holistic humanities of speaking: Franz Boas and the continuing centrality of texts. *International Journal of American Linguistics* 83(1). 41–78.

Evans, Nicholas. 2008. Review of *Essentials of language documentation*. *Language Documentation & Conservation* 2(2). 340–350.

Evans, Nick & Hans-Jürgen Sasse. 2007. Searching for meaning in the Library of Babel: Field semantics and problems of digital archiving. *Language Documentation and Description* 4. 58–99.

Everett, Daniel Leonard. 1983. A língua Pirahã e a teoria da sintaxe: Descrição, perspectivas e teoria. PhD dissertation. Universidade Estadual de Campinas. https://daneverettbooks.com/wp-content/uploads/2014/04/everett_90_A-Lingua-Pirah.pdf (accessed 2021-12-23).

Everett, Daniel L. 1986. Pirahã. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages,* vol. 1, 200–325. Berlin: Mouton de Gruyter.

Everett, Daniel L. 2004. Coherent fieldwork. In Piet van Sterkenburg (ed.), *Linguistics today – Facing a greater challenge*, 141–162. Amsterdam: John Benjamins.

Everett, Daniel L. 2005. Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology* 46(4). 621–646.

Everett, Daniel L. 2006. Cultural constraints on grammar in Pirahã. (Paper presented at the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology.) http://tedlab.mit.edu/tedlab_website/news/everett-recording-2006.wma (accessed 2021-12-23).

Everett, Daniel L. 2007. Cultural constraints on grammar in Pirahã: A reply to Nevins, Pesetsky, and Rodrigues. https://ling.auf.net/lingbuzz/000427 (accessed 2021-12-23).

Everett, Daniel L. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85(2). 405–442.

Everett, Daniel L. 2010. The shrinking Chomskyan corner: A final reply to Nevins, Pesetsky, and Rodrigues. https://ling.auf.net/lingbuzz/000994 (accessed 2021-12-23).

Gawne, Lauren & Andrea L. Berez-Kroeker. 2018. Reflections on reproducible research. In Bradley McDonnell, Andrea L. Berez-Kroeker, & Gary Holton (eds.), *Reflections on language documentation 20 years after Himmelmann 1998*, 22–32 (Language Documentation & Conservation Special Publication 15). Honolulu: University of Hawai'i Press.

Geertz, Clifford. 1973. Thick description: Toward an interpretive theory of culture. In *The interpretation of cultures*, 3–30. New York: Basic Books.

Gibson, Jason, Shaun Penangke Angeles, & Joel Perrurle Liddle. 2019. Deciphering Arrernte archives: The intermingling of textual and living knowledge. In Linda Barwick, Jennifer Green, & Petronella Vaarzon-Morel (eds.), *Archival returns: Central Australia and beyond,* 29–45 (Language Documentation & Conservation Special Publication 18). Honolulu: University of Hawai'i Press.

Good, Jeff. 2018. Ethics in language documentation and revitalization. In Kenneth L. Rehg & Lyle Campbell (eds.), *The Oxford handbook of endangered languages*, 419–440. Oxford: Oxford University Press.

Hammersley, Martyn. 2016. Reflections on the value of ethnographic re-studies: Learning from the past. *International Journal of Social Research Methodology* 19(5). 537–550.

Harrison, K. David. 2005. Ethnographically informed language documentation. *Language Documentation and Description* 3. 22–41.

Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36. 161–195.

Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In Jost Gippert, Nikolaus P. Himmelmann, & Ulrike Mosel (eds.), *Essentials of language documentation*, 1–30. Berlin: Mouton de Gruyter.

Himmelmann, Nikolaus P. 2012. Linguistic data types and the interface between language documentation and description. *Language Documentation & Conservation* 6. 187–207.

Himmelmann, Nikolaus P. 2018. Meeting the transcription challenge. In Bradley McDonnell, Andrea L. Berez-Kroeker, & Gary Holton, (eds.), *Reflections on language documentation 20 years after Himmelmann 1998*, 33–40 (Language Documentation & Conservation Special Publication 15). Honolulu: University of Hawai'i Press.

Holton, Gary. 2014. Mediating language documentation. *Language Documentation and Description* 12. 37–52.

Hymes, Dell. 1974. *Foundations in sociolinguistics: An ethnographic approach*. Philadelphia: University of Pennsylvania Press.

Innes, Pamela. 2010. Ethical problems in archival research: Beyond accessibility. *Language & Communication* 30(3). 198–203.

Innes, Pamela & Erin Debenport (eds.). 2010. Special issue on "Ethical dimensions of language documentation". *Language & Communication* 30(3).

Jung, Dagmar & Nikolaus P. Himmelmann. 2011. Retelling data: Working on transcription. In Geoffrey Haig, Nicole Nau, Stefan Schnell, & Claudia Wegener (eds.), *Documenting Endangered languages: Achievements and perspectives,* 202–220. Berlin: Mouton De Gruyter.

Keane, Webb. 2005. Estrangement, intimacy, and the objects of anthropology. In George Steinmetz (ed.), *The politics of method in the human sciences: Positivism and its epistemological others*, 59–88. Duke: Duke University Press.

Kopytoff, Igor. 1986. The cultural biography of things: Commoditization as process. In Arjun Appadurai (ed.), *The social life of things: Commodities in cultural perspective*, 64–91. Cambridge: Cambridge University Press.

Ladefoged, Peter. 1992. Another view of endangered languages. *Language* 68(4). 809–811.

Latour, Bruno. 1993. *We have never been modern*. Catherine Porter (trans.). Cambridge: Harvard University Press.

Lederman, Rena. 2016. Archiving fieldnotes? Placing 'anthropological records' among plural digital worlds. In Roger Sanjek & Susan W. Tratner (eds.), *eFieldnotes: The makings of anthropology in the digital world*, 251–271. Philadelphia: University of Pennsylvania Press.

Lehmann, Christian. 2001. Language documentation: A program. In Walter Bisang (ed.), *Aspects of typology and universals*, 83–97. Berlin: Akademie Verlag.

Leonard, Wesley Y. & Haley De Korne (eds.). 2017. Special issue on "Reclaiming Languages". *Language Documentation and Description* 14.

Leonard, Wesley Y. & Erin Haynes. 2010. Making 'collaboration' collaborative: An examination of perspectives that frame linguistic field research. *Language Documentation & Conservation* 4. 268–293.

Linguistic Society of America. 2010. Resolution recognizing the scholarly merit of language documentation. https://www.linguisticsociety.org/resource/resolution-recognizing-scholarly-merit-language-documentation (accessed 2021-12-23).

Macri, Martha J. 2010. Language documentation: Whose ethics? In Lenore A. Grenoble & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 37–47. Amsterdam: John Benjamins.

McDonnell, Bradley. 2018. Reflections on linguistic analysis in documentary linguistics. In Bradley McDonnell, Andrea L. Berez-Kroeker, & Gary Holton (eds.), *Reflections on language documentation 20 years after Himmelmann 1998*, 191–200 (Language Documentation & Conservation Special Publication 15). Honolulu: University of Hawai'i Press.

McDonnell, Bradley. Forthcoming. Pragmatic annotation: A method for increasing accessibility of documentation through contextual, ethnographic, and interactional commentary. In Rich Sandoval & Nicholas Jay Williams (eds.), *Interactional Approaches to Language Documentation* (Language Documentation & Conservation Special Publication). Honolulu: University of Hawai'i Press.

McDowell, Nancy. 1991. *The Mundugumor: From the field notes of Margaret Mead and Reo Fortune*. Washington: Smithsonian Institution Press.

Moore, Robert. 2009. From performance to print, and back: Ethnopoetics as social practice in Alice Florendo's corrigenda to "Raccoon and His Grandmother". *Text & Talk* 29(3). 295–324.

Nathan, David. 2013. Access and acessibility at ELAR, a social networking archive for endangered languages documentation. In Mark Turin, Claire Wheeler, & Eleanor Wilkinson (eds.), *Oral literature in the digital age: Archiving orality and connecting with communities*, 21–40. Cambridge: Open Book Publishers.

Nathan, David & Peter K. Austin. 2004. Reconceiving metadata: Language documentation through thick and thin. *Language Documentation and Description* 2. 179–188.

Nettle, Daniel & Suzanne Romaine. 2000. *Vanishing voices: The extinction of the world's languages.* Oxford: Oxford University Press.

Nevins, Andrew, David Pesetsky, & Cilene Rodrigues. 2007. Pirahã exceptionality: A reassessment. https://ling.auf.net/lingbuzz/000411 (accessed 2021-12-23).

Nevins, Andrew, David Pesetsky, & Cilene Rodrigues. 2009. Evidence and argumentation: A reply to Everett (2009). *Language* 85(3). 671–681.

Nevins, M. Eleanor. 2013. *Lessons from Fort Apache: Beyond language endangerment and maintenance*. Chichester: Wiley-Blackwell.

Nevins, M. Eleanor. 2015. "Grow with that, walk with that": Hymes, dialogicality, and text collections. In Paul V. Kroskrity & Anthony K. Webster (eds.), *The legacy of Dell Hymes: Ethnopoetics, narrative inequality, and voice*, 71–107. Bloomington: Indiana University Press.

Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs & Bambi B. Schieffelin (eds.), *Developmental Pragmatics*, 43–72. New York: Academic Press.

O'Meara, Carolyn & Jeff Good. 2010. Ethical issues in legacy language resources. *Language & Communication* 30(3). 162–170.

Powell, Timothy B. 2015. Anthropology of revitalization: Digitizing the American Philosophical Society's Native American collections. In Regna Darnell, Michelle Hamilton, Robert L. A. Hancock, & Joshua Smith (eds.), *The Franz Boas papers, volume 1: Franz Boas as public intellectual – theory, ethnography, activism*, 331–344. Lincoln: University of Nebraska Press.

Rabinow, Paul & William M. Sullivan (eds.). 1987. *Interpretive social science: A second look*. Berkeley: University of California Press.

Reddy, Michael J. 1979. The conduit metaphor – a case of frame conflict in our language about language. In Andrew Ortony (ed.), *Metaphor and thought*, 284–324. Cambridge: Cambridge University Press.

Rice, Keren. 2012. Ethical issues in linguistic fieldwork. In Nicholas Thieberger (ed.), *The Oxford handbook of linguistic fieldwork*, 407–429. Oxford: Oxford University Press.

Sanjek, Roger (ed.). 1990. *Fieldnotes: The makings of anthropology*. Ithaca: Cornell University Press.

Schmidt, Thomas & Jasmine Bennör. 2008. Rescuing legacy data. *Language Documentation & Conservation* 2(1). 109–129.

Schwartz, Saul. 2019. The Boas plan: A view from the margins. In Regna Darnell & Frederic W. Gleach (eds.), *Disruptive voices and the singularity of histories*, 281–318 (Histories of Anthropology Annual 13). Lincoln: University of Nebraska Press.

Schwartz-Shea, Peregrine. 2014. Interpretive social science. In Michael T. Gibbons (ed.), *The encyclopedia of political thought*. New York: Wiley-Blackwell. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118474396.wbept0533 (accessed 2021-12-23).

Seidel, Frank. 2016. Documentary linguistics: A language philology of the 21st century. *Language Documentation and Description* 13. 23–63.

Silverman, Sydel & Nancy Parezo (eds.). 1995. *Preserving the anthropological record*. New York: Wenner-Gren Foundation for Anthropological Research.

Silverstein, Michael. 1996. The secret life of texts. In Michael Silverstein & Greg Urban (eds.), *Natural histories of discourse*, 81–105. Chicago: University of Chicago Press.

Silverstein, Michael. 2015. From Baffin Island to Boasian induction: How anthropology and linguistics got into their interlinear groove. In Regna Darnell, Michelle Hamilton, Robert L. A. Hancock, & Joshua Smith (eds.), *The Franz Boas papers, volume 1: Franz Boas as public intellectual – theory, ethnography, activism*, 83–127. Lincoln: University of Nebraska Press.

Silverstein, Michael. 2017. The fieldwork encounter and the colonized voice of indigeneity. *Representations* 137. 23–43.

Stebbins, Tonya N. 2012. On being a linguist and doing linguistics: Negotiating ideology through performativity. *Language Documentation & Conservation* 6. 292–317.

Steinmetz, George (ed.). 2005. *The politics of method in the human sciences: Positivism and its epistemological others*. Durham: Duke University Press.

Stocking, George W. 1992. *The ethnographer's magic and other essays in the history of anthropology*. Madison: University of Wisconsin Press.

Thieberger, Nicholas (ed.). 1995. *Paper and talk: A manual for reconstituting materials in Australian Indigenous languages from historical sources.* Canberra: Aboriginal Studies Press.

Thieberger, Nicholas & Andrea L. Berez 2012. Linguistic data management. In Nicholas Thieberger (ed.), *The Oxford handbook of linguistic Fieldwork*, 90–118. Oxford: Oxford University Press.

Thieberger, Nicholas & Michel Jacobson. 2010. Sharing data in small and endangered languages: Cataloging and metadata, formats, and encodings. In Lenore A. Grenoble & N. Louanna Furbee (eds.), *Language documentation: Practice and values*, 147–158. Amsterdam: John Benjamins.

Tobin, Joseph. 1990. The HRAF as radical text? *Cultural Anthropology* 5(4). 473–487.

Trouillot, Michel-Rolph. 1995. *Silencing the past: Power and the production of history.* Boston: Beacon Press.

Voegelin, C. F. 1952. The Boas plan for the presentation of American Indian languages. *Proceedings of the American Philosophical Society* 96(4). 439–451.

Wasson, Christina, Melanie Medina, Miyoung Chong, Brittany LeMay, Emma Nalin, & Kenneth Saintonge. 2018. Designing for diverse user groups: Case study of a language archive. *Journal of Business Anthropology* 7(2). 235–267.

Whalen, D. H. & Joyce McDonough. 2018. Under-researched languages: Phonetic results from language archives. In William F. Katz & Peter F. Assmann (eds.), *The Routledge handbook of phonetics*, 51–71. London: Routledge.

Wierzbicka, Anna. 2005. Comment. *Current Anthropology* 46(4). 641.

Woodbury, Anthony C. 2003. Defining documentary linguistics. *Language Documentation and Description* 1. 35–51.

Woodbury, Anthony C. 2007. On thick translation in linguistic documentation. *Language Documentation and Description* 4. 120–135.

Woodbury, Anthony C. 2011. Language documentation. In Peter K. Austin & Julia Sallabank (eds.), *The Cambridge handbook of endangered languages*, 159–211. Cambridge: Cambridge University Press.

Woodbury, Anthony C. 2014. Archives and audiences: Toward making endangered language documentations people can read, use, understand, and admire. *Language Documentation and Description* 12. 19–36.

Yanow, Dvora & Peregrine Schwartz-Shea. 2014. Wherefore 'interpretive': An introduction. In Dvora Yanow & Peregrine Schwartz-Shea (eds.), *Interpretation and method: Empirical research methods and the interpretive turn*, xiii–xxxi. 2nd edn. London: Routledge.

Yearley, Steven. 1988. The dictates of method and policy: Interpretational structures in the representation of scientific work. In Michael Lynch & Steve Woolgar (eds.), *Representation in scientific practice*, 337–355. Cambridge: MIT Press.