

Transcription as an iterative and interpretive practice: Learning from the documentation of connected speech in Apsáalooke (Crow)

EDWIN KO*, Yale University, US, eddersko@gmail.com

JEM BURCH, Yale University, US, jem.burch@yale.edu

*Corresponding author

ABSTRACT

In this paper we examine a key aspect of documentary linguistic practice by reflecting on our approach to the transcription of recorded narratives in Apsáalooke (Crow), a Siouan language primarily spoken in Montana, USA. We discuss the creation of transcripts that serve multiple audiences and highlight transcription as an iterative and evolving process that can sometimes give rise to multiple interpretations. In particular, we describe how a single audio recording may yield several valid interpretations, not only by different individuals but also the same person on separate occasions. This finding underscores how factors such as collaboration, timing, and setting influence the resulting transcript. Since transcripts play a crucial role in linguistic research and the development of community-oriented resources, we advocate for greater self-reflection in transcription practices. We conclude by considering the implications for transcription and digital archiving practices and provide practical recommendations for researchers engaged in language documentation.

Keywords: transcription; language documentation; Crow

1 INTRODUCTION

Since the advent of audio recording technology and the push towards documenting endangered languages, there has been an immense effort to document, describe, and preserve signed and spoken languages worldwide. However, literature on language documentation and description has rarely addressed how researchers should go about deciding what to transcribe and what the process of transcription should look like (Himmelman 2018). Yet, issues relating to the practice of transcription have been discussed at length in other subfields of linguistics, such as conversation analysis and linguistic anthropology, where there has been a call for greater reflection on how one goes about creating a transcript of spoken discourse (e.g., Bucholtz 2000; Mondada 2007; Ochs 1979).

By recording the speech/signs and cultural practices of an individual or a group, it is possible to preserve their voice, likeness, and activities for posterity. Yet even with all the benefits of modern technologies, recording spoken discourse strips away contextual information from the original speech event, including the physical space, gestures, and audience. Additionally, recording transforms the discourse into a mobile, portable object that is circulable and mutable. In other words, the evanescent and transient nature of sound is concretized—or as linguistic anthropologists say, entextualized—when it becomes a physical record. Through inscription discourse gains new contextual meanings whenever it is revisited, a process often referred to as recontextualization.

Although the literature on entextualization and recontextualization is substantial (see Bauman & Briggs 1990; Silverstein & Urban 1996, and references therein), the same cannot be said for the implications of transcription practices within language documentation. Himmelman (2018: 34, fn.2), references only Crowley's 2007 *Field linguistics: A beginner's guide*, which assumes that the reader uses a cassette recorder.¹ In the section on text transcription, Crowley (2007: 138) writes, “[y]ou will need to transcribe exactly what was said on tape, as well as its meaning”. It bears asking what Crowley meant by “exactly what was said”, and whether obtaining an objective transcript is truly achievable (see Bucholtz 2000, who argues powerfully that it is not).²

As speech/signs are entextualized and recontextualized, multiple interpretations naturally arise (Austin 2017; Dobrin & Schwartz 2021; Haviland 1996; Urban 1996; Weber 2021). As we describe in this paper, a single audio recording may engender distinct interpretations not only by different individuals but also the same person on separate occasions. This finding indicates that the choice of who one works with, and when and where that work occurs, etc., can shape each working transcript. Furthermore, because new insights may surface every time a text is revisited, we suggest that transcription ideally should progress iteratively to reflect the evolution of the analysis. Since transcripts play a crucial role both in linguistic research and in the development of community-oriented resources, we encourage language documentarians to turn a self-reflective gaze on their transcription process. We offer several recommendations for how to go about

1 Other writings that discuss the transcription process before the publication of Himmelman (2018) include, e.g., Bower (2015: 136), Chelliah and de Reuse (2011: 431–442), and Foley (2002: 136–137). More recently, Epps et al. (2023: 189–192), also discuss approaches to transcribing connected speech.

2 In this paper, we draw a distinction between ‘transcription’ and ‘transcripts’. We use ‘transcription’ to refer to the process of developing a written representation of another medium—in this case, digital recordings—in an orthography. We use ‘transcripts’ to refer to the written representations obtained through the process of transcription, which may include supplementary analytical components such as corresponding morpheme-by-morpheme breakdowns, interlinear glosses, and translations. The format and level of detail in a transcript can vary depending on the intended audience. Throughout this paper, when we refer to ‘transcripts’, we focus on the top line that displays the orthographic representation of the language, as this has direct consequences on subsequent analyses.

this in the conclusion of this paper, where we also discuss the implications of the transcription process for transcription and digital archiving practices, and their effects on researchers and linguists, who almost always rely on transcripts to produce language revitalization materials and to study language structure and use.

Our proposal that transcription is an iterative process is not novel. In an 1889 paper, Franz Boas comments that his transcription of Inuktitut improved over time as his ability to discern the sounds of the language increased. Working on the rhetorical structure of narratives, Virginia Hymes (1987:68, emphasis hers) explicitly remarks, “what is involved is a patient working back and forth between content and form [...] until gradually you arrive at an analysis that seems best *at that stage of your knowledge of the narrative tradition and the particular narrator*”. These examples underscore how the evolution of the researcher’s perspectives can necessitate the revision and re-revision of proposed analyses (see also Evans and Sasse 2007; Mondada 2007; Seidel 2016).

The idea that a single recording is subject to more than one interpretation is also not new. Researchers and community collaborators alike make conscious and unconscious decisions while transcribing, and transcribers themselves may also filter the content of a recording. For example, Dobrin (2021) compares transcriptions of texts in Bukiyip, a language of Papua New Guinea, with the original audio and identifies several discrepancies. In particular, the transcriber, a native speaker of the language, made decisions to purge repetitions and ‘purify’ the recording by removing Tok Pisin words, retranslating those portions into Bukiyip (see also Chelliah 2001; Haviland 1996; Jung & Himmelmann 2011; Urban 1996; Weber 2021).³

Given the number of potential transcripts, which of them should be maintained for future users? Although Jung and Himmelmann (2011: 219) state that “[t]here is no straightforward and easy answer to this question”, we propose that transcripts of *all* interpretations should be maintained and archived so long as it is ethical to do so (see Evans and Sasse 2007; Innes 2010). By preserving all versions of transcripts, we maintain each individual’s distinct but equally valid interpretation of the same text, which may potentially provide insights that would otherwise be lost. The structure of this paper is as follows. In Section 2, we describe the data analyzed, the project for which the data were collected, and the transcription process. In Section 3, we discuss the types of transcripts we use, how they reflect the goals of our project, and the challenges we encountered during the transcription process. In Section 4, we explore several types of variation found in the transcripts, from variation in the interpretation of the narratives to variation in forms in Apsáalooke (Crow) across different speakers. In Section 5, we examine the implications of this variation for transcription and archival practices. Finally, Section 6 offers concluding remarks, along with a set of considerations and lessons learned for other field researchers engaged in transcribing spoken discourse.

2 DATA AND THE TRANSCRIPTION PROCESS

The data analyzed in this paper derives from the first author’s field recordings, which are housed in the California Language Archive of the Survey of California and Other Indian Languages maintained by the University of California, Berkeley (Alden et al., n.d.). Copies of these recordings have also been shared with

3 We experienced a similar issue when working directly with speech communities undergoing rapid language shift to English. When informed that these recordings would be archived and preserved for future generations, speakers adapted their speech to be viewed as more ‘authentic’, omitting instances of code-switching and coining terms they did not know or remember. In doing so, these individuals presented themselves and perhaps their larger community as unaffected by the dominant society. In contrast, in many settings without such conditions, Apsáalooke-English code-switching was a common practice.

the Little Big Horn College, the tribal college located in Crow Agency, the headquarters of the Crow Indian Reservation in Montana. Since 2018, the first author has been working with community collaborators to record oral histories pertaining to places on and around the reservation, along with personal reminiscences about changes to the reservation and the way of life, the contemporary language situation, and aspects of Apsáalooke identity and culture. As the Apsáalooke language is currently endangered, our project aims to address this issue in three ways. The first goal is to preserve Apsáalooke history, culture, and social structure as it exists in the early 21st century. The second is to create an online interactive storytelling map using Esri's *ArcGIS StoryMap*, featuring the recorded narratives and accompanying transcripts. This platform enables community members to freely access and engage with the materials, thereby supporting future language revitalization and reclamation efforts. Finally, this project aims to document connected speech for more advanced learners of the language, who may be interested in how it is used in informal contexts, and to inform future work on understudied aspects of Apsáalooke language and culture.

To clarify the contributions of each author in the article, we use the pronoun 'I' to refer to efforts conducted by Edwin Ko, who has spent many years collaborating with community members on the Crow Indian Reservation to conduct fieldwork on the Apsáalooke language. The use of the first-person pronoun is not intended to minimize the contributions of Jem Burch, an undergraduate Yale linguistics student who served as a research assistant on the project. In fact, Jem Burch played a crucial role in the research, actively participating in the data collection during the summer of 2024, and significantly contributing to the analysis, literature review, drafting, and revising of the paper. The pronoun 'we' is employed when referring to collaborative efforts undertaken by both authors, emphasizing our joint responsibility and collaboration in the study's development.

The project to record oral narratives emerged from sustained dialogue with community members, during which I carefully listened to their concerns. Although there were no explicit expectations regarding my role in language revitalization, community members frequently voiced anxieties not only about the future of the language but also about broader transformations in social structure, cultural practices, and relationships to the land, of which one of the most visible consequences is widespread urban decay, characterized by abandoned or dilapidated buildings and increasing family fragmentation. In addition, the project complements ongoing efforts at the tribal college, including a National Endowment for the Humanities-funded initiative to document oral histories from the post-World War II era. Although no formal approval process was established, the project was met with positive feedback from all community members with whom I engaged. Building on relationships formed during my initial visit to the community in 2016, when I collaborated with the then-Head of Education of the Crow Tribe, I recruited volunteers willing to share their stories. Individual sessions were conducted primarily at the Little Big Horn College library and the public library in the nearby border town of Hardin, or, in the case of place-based oral histories, at the relevant sites.

During a typical recording session, we would begin by inviting the community collaborator to introduce themselves and then ask them a range of open-ended questions about their perspectives on the Apsáalooke language and culture. These questions addressed topics such as the current state of the language, generational differences in how Apsáalooke is spoken, and what it means to be Apsáalooke when speaking to future generations. For the place-based oral histories, collaborators were asked to recall personal experiences and memories directly associated with specific extant and historical sites. In each case, collaborators were encouraged to answer in both English and Apsáalooke, and all responses were recorded. Following some of the sessions, we produced a transcript of the recording.

To illustrate our transcription process, we describe our transcription of a narrative contributed by community collaborator Theresa Sends Part Home, a woman born in the late 1950s, who had fond memories growing up in Lodge Grass, one of the towns on the reservation. I completed a first pass of the transcriptions using the time-aligned annotation tool ELAN (Wittenburg et al. 2006). During the COVID-19 pandemic, I collaborated remotely with Felice Big Day, an Apsáalooke speaker born in the late 1970s with whom I had previously worked, to transcribe the narratives. This interpretation session was conducted over the phone while I was in California and Felice was in Montana. To transcribe the recording, the entire audio of Theresa's narrative was played in Apsáalooke. Then, Felice repeated and translated each word. Whereas the recording of a single narrative typically took place in a single session, the transcription process would often span several sessions. After the recordings had been transcribed in ELAN, I heavily revised them before transferring them to LaTeX to be formatted into interlinear text. The transcriptions were written in the "standard" Apsáalooke orthography, which omits aspects of rapid, informal speech. Recordings of all interpretation sessions were deposited at the California Language Archive before they were shared with the Little Big Horn College.

When it was safer to visit the reservation, I collaborated with Felice in person to reinterpret the same stories provided by Theresa. This time, rather than editing the ELAN file, I transcribed each word in my field notebook alongside the corresponding translations. After studying the recordings and my field notes, I made corrections to the original LaTeX file. After repeatedly listening to both Theresa's recorded oral narrative and Felice's recorded interpretation, I resolved any remaining issues with Felice. Once the transcriptions and glosses were finalized, they were transferred back to ELAN to be archived.

While this transcription process is largely consistent with that of other language documentarians (see, e.g., Chelliah and de Reuse 2011: 431–442), it differs from the practices used in other disciplines. As conversation analysts Hepburn and Bolden (2013: 57) state, "Conversation analytic transcripts need to be detailed enough to facilitate the analyst's quest to discover and describe orderly practices of social action in interaction". In the field of ethnopoeitics, scholars aim to "[arrange] the transcript of an event of oral narration so as to reflect or recuperate the true rhetorical architecture of denotational text" (Moore 2013: 14). In these disciplines, transcription plays a fundamental role in allowing researchers to identify patterns that serve the specific communicative purpose being studied.

In contrast, the purpose of language documentation, as Woodbury (2010: 159) writes, is the "creation, annotation, preservation and dissemination of transparent records of a language". That said, transparency is easier to aim for than to achieve. Although field notebooks may preserve some version of a given text, various challenges, such as idiosyncratic notation or missing contextual information, can reduce their transparency and accessibility for future users (see, e.g., Goddard 1976; Austin 2017; Begay et al. 2021). In many cases, archives contain only a single digital transcript, omitting any record of the iterative revisions that shaped the final version. Furthermore, even archival recordings that feature the voice of only a single speaker, as many of our Apsáalooke recordings do, cannot be treated as entirely monologic. As numerous scholars have argued, interviews are co-constructed by all participants, including the researcher, and are shaped by cultural expectations and the dynamics of the interaction itself (Briggs 1986; Haviland 1996; De Fina and Perrino 2011; Dobrin 2012).⁴

4 In fact, some consultants noted the awkwardness of speaking to researchers not fluent in the language, and often found it difficult to conjure up an imagined audience proficient in the language who could understand what they were saying (see Bower 2015: 134).

At every stage of the transcription process, some information inevitably gets lost. Miscommunication may occur—a researcher might mistake a word that underwent syncope for another word, or misinterpret false starts and hesitation markers as part of the previous or following words (see Jung and Himmelmann 2011: 211). After consulting with speakers, the researcher may discover their errors and correct them, but the revision process remains hidden if only the final version of the transcript is archived. It is human nature to be unwilling to reveal past mistakes, but presenting the evolution of transcripts can be insightful for others. Furthermore, if a researcher working with multiple speakers receives varying interpretations, but publishes only the interpretation of a single speaker, it is important to include in the metadata an explanation for why that particular interpretation was selected.

Every stage of the transcription process builds upon a previous stage. Even before listening to a recording, each user will contribute their own prior knowledge, lived experiences, and biases—what Fairclough (1989) and Bell (2011) call the “members’ resources” and “pre-view”, respectively. Therefore, we advocate that language documentarians archive multiple versions of each transcript. We also suggest that they provide relevant metadata, such as their research agenda and biography, to contextualize their collection for future users (Austin 2017; Dobrin & Schwartz 2021; Nathan & Austin 2004; Innes 2010; Woodbury 2014). For full disclosure, we have not archived all earlier versions of our transcripts; earlier drafts produced solely by me were replaced with later interpretations offered by speakers. It was only when I conceived the idea that laid the groundwork for this paper that we began preserving these initial drafts.

3 TRANSCRIPTS

In this section, we describe our approach to transcribing Apsáalooke texts (Alden et al., n.d.) based on our goals—to provide transcripts for future researchers and language-learners—and reflect on its drawbacks.^{5,6} For this, we find Bucholtz’s (2000: 1461–1463) distinction between *naturalized* and *denaturalized* transcripts particularly useful. Naturalized transcripts closely adhere to the standard written conventions, whereas denaturalized transcripts preserve aspects of the signed or spoken language, which may include incorporating paralinguistic features of signed and spoken language, such as pauses and intonation, into the transcripts. In some sense, naturalized transcripts are highly edited versions of denaturalized ones. However, as noted by Bucholtz (2000), both are extremes on a continuum, and neither is optimal. We have chosen to represent our texts in these two styles. Transcripts presented in a naturalized style expose learners of Apsáalooke to a standard written form without variation, fast-speech phenomena, etc.; transcripts presented in a denaturalized style allow advanced learners to study informal speech and provide data for future analysis of understudied areas of Apsáalooke linguistics.

The naturalized transcripts are written in the standard Apsáalooke orthography, which maintains information about pitch accent. However, these transcripts are not totally naturalized in the same way Bucholtz originally presented them, as we also preserve discourse features, such as hesitation markers, in order to highlight that the speech was performed orally rather than read from a script. The denaturalized transcripts preserve some fast-speech phenomena and other discourse features that would otherwise

5 In the literature, the term ‘texts’ can denote both written and spoken discourse. In this paper, the use of this term refers to spoken discourse by default. Whenever we use ‘texts’ to mean written discourse, we will make this explicit.

6 Due to space, we omit discussion on how transcripts should be formally presented. We also exclude detailed discussion of glossing conventions and translations of the texts themselves.

be omitted, including repetitions, false starts, self-corrections, and hesitation particles. However, we exclude intonation, duration of pauses, audible breathing, laughter, and other features that might be important to transcripts produced for conversation analysis and ethnopoetics. Our decision to exclude these features is not born out of a belief that such discursive features are unimportant, but as Ochs (1979: 44) remarks, “transcription is a selective process” that reflects one’s own goals. Being more selective about what we transcribe also reduces barriers of access to the transcripts, as including these additional annotations would decrease readability, especially for those focused on language revitalization.

Excluding paralinguistic features thus highlights the project’s overall purpose, which is primarily concerned with the content of the oral history and its importance to the Apsáalooke community. However, this comes at the expense of showcasing the speaker’s voice, that is, the “freedom to have one’s voice heard, freedom to develop a voice worth hearing” (Hymes 2004: 64; see also Dobrin 2012, 2021; Webster 2013). The presence of the interviewer is also omitted, based on the assumption that users are primarily interested in the voices of community members. Yet by presenting these transcripts as monologic, we inevitably obscure aspects of the participant framework—namely, the configuration of roles, statuses, and alignments occupied by each participant, including the researcher, in relation to others and to the unfolding interaction at a particular moment in time (Goffman 1981).

However, by offering two distinct transcripts, we are able to both cater to multiple audiences and escape the difficulty of choosing between a transcript that is overly technical and one that is overly simplistic (Jung and Himmelmann 2011: 207). Below, we provide examples of naturalized and denaturalized transcription styles from our collection, using excerpts from narratives offered by Theresa Sends Part Home as a case study.

3.1 Naturalized transcript

An example of a naturalized transcript, in which Theresa Sends Part Home provides a brief 30-second introduction, is given in (1).⁷

(1) Naturalized transcript (transcribed by Edwin Ko with the assistance of Felice Big Day on 1 November, 2022):

- a. Ba-láshe Theresa Sends Part Home huu=k.
1POSS-name Theresa Sends Part Home it.is.said=DECL
‘My name is Theresa Sends Part Home.’
- b. Hem, uh, bíilukaa-ilaa=u ba-láshe Ak=chiwakii-sitchi=sh huu=sh.
FILL FILL Apsáalooke-speak=PL 1POSS-name REL=pray-like=DEF. say.PL=DEF
‘And, uh, my Apsáalooke name is *Akchiwakisitchish* (One Who Likes To Pray).’
(Theresa Sends Part Home; Theresa_080218_000.wav: 1:37–1:47)

The transcript presented in (1) conforms to standard Apsáalooke orthographic conventions, making it more similar to a naturalized transcript than a denaturalized one. As is typical for transforming oral

⁷ The abbreviations used in the glosses are as follows: 1 first person; 2 second person; A active; AGT agentive; AUG augmentative; B stative; BEN benefactive; CAUS causative; CONJ conjunction; DECL declarative; DEF definite; DESID desiderative; DIMIN diminutive; DIR direct; DS different-subject; FOC focus; FILL filler; FS false start; HAB habitual; INDEF indefinite; INDIR indirect; INSTR instrumental; JUNCT juncture/linker; MOD modal; NEG negative; NONAGT non-agentive; OBJ object; PL plural; POSS possessive; PRO pronominal; PST past; REL relativizer; SS same-subject; TEMP temporal.

English discourse into prose, commas in the text primarily delineate fillers; they do not denote properties like rising intonation as they might in transcripts by conversation analysts.⁸

Like other non-referential aspects of the language, hesitation markers in Apsáalooke are poorly documented. There are several reasons to believe that *hem* is one such marker. To our knowledge, *hem* is excluded from all written materials in or about Apsáalooke produced by the community, missionaries, and previous linguists, suggesting that it does not encode any referential meaning and/or alter the propositional content of the sentence. However, this marker can be found in abundance in our recordings of oral narratives, appearing in the speech of most participants. While it generally occurs clause-initially, as in (1b), it may also occur within phrases and typically co-occurs with English fillers (e.g. *uh*, *um*). Furthermore, Felice offered that “it’s kind of like saying, when you’re talking English, you’re saying *um*”, and Randolph Graczyk (personal communication, 2024-06-09), a former pastor on the reservation and senior Apsáalooke scholar, commented that *hem* is “simply a conversational filler”.

In addition to inserting commas, we chose to capitalize the first letter of each sentence and conclude each sentence with a period, following the English orthographic conventions reappropriated into Apsáalooke. However, determining where the sentence begins and ends is no easy task (Himmelmann 2006; Chelliah and de Reuse 2011: 437–438). In most cases, there are (para)linguistic cues that can indicate where the boundaries of a sentence are located, such as pauses, prosody, discourse markers, and verb-final morphology. For instance, *hem* is typically found clause-initially, while the declarative clitic =*k* and other speech-act markers are often found sentence-finally. In other cases, intuitions provided by speakers and educated guesses by us are incorporated into the analysis.

3.2 Denaturalized transcript

Unlike naturalized transcripts, denaturalized transcripts aim to preserve more features of spoken discourse. Example (2) shows the denaturalized version of the same naturalized transcript seen in (1).

(2) Denaturalized transcript:

```
ba-láshe      Theresa Sends Part Home huu=k          hem  b-  uh
1POSS-name Theresa Sends Part Home it.is.said=DECL  FILL FS  FILL

Bíilukaa-ilaa=u      ba-láshe      Ak=chiwakíi-sitchi=sh huu=sh
Apsáalooke-speak=PL 1POSS-name REL=pray-like=DEF      say.PL=DEF
‘My name is Theresa Sends Part Home. And, uh, my A- Apsáalooke name is Akchiwakíisitchish (One
Who Likes To Pray).’
(Theresa Sends Part Home; Theresa_080218_000.wav: 1:37–1:47)
```

In (2), we include false starts (indicated by hyphens), but we do not use punctuation to delineate fillers, false starts, or sentential boundaries. This choice allows the transcript to serve as a ‘baseline’ for future investigation into aspects of the Apsáalooke language. Depending on the transcription conventions employed and the theoretical leanings of the linguist, commas and periods may be interpreted as rising or falling intonation, or as other prosodic features. To avoid such confusion, we do not employ the written conventions of the naturalized transcript. However, we retain capitalization of proper nouns, such as *Theresa Sends Part Home*.

8 The terms hesitation markers and fillers are used interchangeably throughout this paper.

While the passage in (2) does not contain any fast-speech properties, longer recordings often do. An excerpt from a longer transcript of Theresa's reminiscences about Cozy Corner, a cafe in Lodge Grass that burned down in 1992, is shown in (3). The naturalized and denaturalized transcripts are provided in (3a) and (3b), respectively. In the transcripts below, sounds enclosed in parentheses indicate elided segments. The bolded portions illustrate differences in how fast-speech phenomena are represented across the two transcription styles.

(3) Transcribed by Edwin Ko with the assistance of Felice Big Day on 4 November, 2022:

a. Naturalized transcription:

Bassée TV aw-ákaa=**wia=waa**=i=lu=t bassée TV, um, uh,
 long.ago TV 1A-see=**DESID=1A**=HAB=PL=TEMP long.ago TV FILL FILL
 stations, uh, stations kooté **dúup**=kaati=i=k.
 stations FILL stations like.that two=DIMIN=HAB=DECL

b. Denaturalized transcription:

bassée TV (a)w-ákaa=**waai**=(i)=lu=t bassée TV um uh stations
 long.ago TV 1A-see=**1A.DESID**=HAB=PL=TEMP long.ago TV FILL FILL stations
 uh stations kooté **lúup**=kaati=i=k
 FILL stations like.that two=DIMIN=HAB=DECL
 'Long ago, when we used to try to watch TV, there would only be two TV stations.'
 (Theresa Sends Part Home; Theresa_080218_006.wav: 3:17–3:23)

In fast speech, the morphemes *-wia* and *-waa* seen in (3a) are sometimes reduced to a portmanteau form *-waai*, as shown in (3b). Additionally, when the previous word ends in a vowel or nasal, a word-initial segment *d* can surface as *l* or *n*, respectively.⁹ This process is seen in (3b), where the word *dúupkaatiik* in (3a) is realized as *lúupkaatiik* after the form *kooté*.

3.3 Reflections on the transcription process

As Bucholtz (2000) argues, all transcripts carry political and social implications, whether or not such associated connotations are intentional, and ours are no exception. Because these transcripts were created without input from Theresa Sends Part Home, who was unfortunately unavailable, we relied on interpretations by another community collaborator that may not reflect what Theresa had originally intended or how she wanted to be represented. The Apsáalooke orthography employed in our transcripts is used by the Bilingual Materials Development Center, the Wycliffe Bible translators, the Crow Language Consortium, and is actively taught at the Little Bighorn College and the Crow Summer Institute. But as many fieldworkers know well, there is often debate over which (and whose) orthographic conventions to adopt (see Cahill & Rice 2014). Not everyone in the community accepts the “standard” Apsáalooke orthography. In fact, I worked closely with a speaker of Apsáalooke who preferred an orthography she had invented *de novo* based on her idiolect for the development of some language-learning materials.

Questions also arise about variation in form. Does elision of segments constitute dialectal variation? Is it merely an effect of casual speech or indicative of a different word altogether? For example, the

9 The segment *d* surfaces word-initially and adjacent to obstruents, whereas *l* surfaces intervocalically.

word ‘sweat lodge’ exhibits inter-speaker variation: some speakers use the form *awúsuaa*, as Theresa does in her narratives, while others use *alawúsuaa*. According to Felice Big Day, the form *awúsuaa* originated from *alawúsuaa*, which features the relativizer *ala-*. Yet when we discussed the same pair of words with a faculty member at the Little Big Horn College, he mentioned that these variants may suggest dialectal variation.¹⁰ A reviewer notes that, in their experience, *awúsuaa* is more commonly used in spoken discourse than *alawúsuaa*. They also highlight that the entry for *alawúsuaa* in the Crow Dictionary specifies that the term refers not to the physical sweat lodge itself but to “the sweat lodge as a ceremonial property” (Gordon & Graczyk, n.d.: 47). Given the absence of in-depth sociolinguistic studies on variation in Apsáalooke and in light of the language’s ongoing shift—during which subtle semantic distinctions, particularly those involving culturally significant terms, may erode (Mithun 1989: 248)—it remains unclear how best to address such variation in our transcripts.

4 VARYING INTERPRETATIONS

The following four case studies illustrate the variability that emerges from our transcription process. The first compares interpretations conducted remotely and in-person with the same community collaborator, Felice Big Day. The second highlights two distinct interpretations of the same text also by Felice, received just five days apart. The third examines how two different collaborators produced divergent interpretations of the same recording. The final case focuses on interpretations of the same text from two different community collaborators, one of whom was the original contributor to that text. It should be noted that the differences observed in our transcripts, including those presented in this paper, primarily concern discourse markers or variants, rather than elements that would alter the core semantic content of the utterance, which is consistent with what has been observed by previous scholars (e.g. Jung and Himmelmann 2011; Webster 2013; see also Kowal and O’Connell 2014). In what follows, all transcripts are presented in a denaturalized style.

4.1 Remote vs. in-person

After I transcribed Theresa’s oral histories, I worked with Felice to interpret the recordings. Interpreting oral histories over the phone and on Zoom—a video conferencing platform—poses many challenges. Because the transcription session took place remotely, my laptop speakers filtered the acoustic signal before it was transmitted over the phone. Not only was audio quality poor on both ends, but several seconds of delay often necessitated frequent hesitations and long pauses between turns. These factors disrupted the natural rhythm of conversation and introduced friction into the interpretive process.

In fact, mishearing is a common occurrence when relying on digital communication technologies, particularly under poor signal conditions (Boland et al. 2022). We noted this in the remote interpretation sessions. For instance, when Felice uttered *ítchissaache* ‘to favor someone or something’, I mistook it as *ítchitaache* and repeated the illicit Apsáalooke form. Felice, in turn, misheard *ítchitaache* as *ítchissaache*, unintentionally confirming that the illicit form was indeed the same word she had uttered just moments before.

10 Felice Big Day, who is from Pryor, says *ishé* rather than *ashé* ‘house; clan; tribe’ and *alúke* instead of *ilúke* ‘meat’. When asked about *ilúke*, she commented that this form is “from over here [Crow Agency]”. A reviewer provides the following commentary: “There are lots of cases where *a* and *i* (unaccented) vary in words. It is common to hear *bis-* instead of *bas-* (possessive pronominal prefixes) e.g., *bisbilaxpáake* for *basbilaxpáake* [‘my people’]. This appears to be a regularization of the paradigm. There is also variation between *a* and *u*, e.g., *xalússhi* vs. *xulússhi* [‘run’], *biilápaache* vs. *biilúpaache* [‘friend’]. These variations need to be further investigated to see if there are really dialectal variants among districts”.

It was not until 2022, after the periodic surges in positive cases of COVID-19 had steadily declined, that it was safer to visit community collaborators on the reservation. At that time, I worked in-person with Felice to translate and interpret the same recordings that both of us had gone over during the pandemic. The following example features two transcripts of another one of Theresa’s recordings about Cozy Corner. Example (4a) is from the transcript produced during the remote session, while (4b) is from the in-person session two years later. (This is the same stretch of talk that was presented above in (3).)

- (4) a. Transcribed by Edwin Ko with the assistance of Felice Big Day on 7 February, 2020:

bassée TV (a)w-ákaa=wai=**w=uu**=t bassée TV um uh
 long.ago TV 1A-see=1A.DESID=**1A=MOD.PL**=TEMP long.ago TV FILL FILL
 stations kooté lúup=kaati=i=**lu**=k
 stations like.that two=DIMIN=HAB=**PL**=DECL
 ‘Long ago, when we **would** want to watch TV, there would only be two TV stations.’

- b. Transcribed by Edwin Ko with the assistance of Felice Big Day on 4 November, 2022:

bassée TV (a)w-ákaa=wai=**(i)=lu**=t bassée TV um uh
 long.ago TV 1A-see=1A.DESID=**HAB=PL**=TEMP long.ago TV FILL FILL
 stations kooté lúup=kaati=i=k
 stations like.that two=DIMIN=HAB=DECL
 ‘Long ago, when we **used to** try to watch TV, there would only be two TV stations.’
 (Theresa Sends Part Home; Theresa_080218_006.wav: 3:17–3:23)

There are two main differences in how Felice heard and repeated the utterances in the recordings, leading to discrepancies in the transcription of her interpretations in (4). First, the morpheme *-wu*, which functions as a first-person plural future-oriented modal (cf. Graczyk 2007: 116, Ex.56), is heard in (4a). In contrast, *-ilu*, the plural habitual, is heard in (4b). These different interpretations led to slight differences in the translations. The morphemes in question were translated as ‘would’ in (4a) and ‘used to’ in (4b). The latter is used to express that watching TV was a routine or habitual practice, whereas the former does not necessarily carry the same implication.

The second difference involves the final word in the sentence, a stative verb. Since the predicate *dúupe* ‘to be two’ semantically entails more than one, plural marking is optional. In (4a), the plural marker *-lu* occurs on the verb, but in (4b), it is absent; instead, *-lu* appears only on the verb in the first clause. Although both sentences in Example (4) are grammatically well-formed and felicitous in the context of the narrative, the discrepancies in Felice’s understanding of the recording and my subsequent corresponding transcripts illustrate that the conditions under which a transcription is produced can influence how one hears and interprets a given narrative.

4.2 Same community collaborator

We next investigate discrepancies that can arise when working with the same collaborator to interpret the same text in person a few days apart. Although we found no major differences, some minor changes carry implications for analyzing the discursive structure of the language. The transcripts in (5) come from a narrative by Theresa, speaking to an imagined future audience, about what being Apsáalooke means to her.

- (5) a. Transcribed by Edwin Ko with the assistance of Felice Big Day on 1 November, 2022:

helak um b-akúpee=sh bah-kuxsá=k bas-bilaxpáakee=sh kon
 so.then FILL 1POSS-sibling=DEF 1A-help=SS 1POSS-people=DEF PRO
 mah-kuxssá=k uh ammaa=ítchee=sh xaxúa **ko**
 1A-help=SS FILL REL=good=DEF all **PRO.NONAGT/FOC**
 waaátchim-ma=htaa=(a)k bah-chiwakáa=(a)=wa=ku=k
 be.fortunate-1A=INDIR.CAUS=SS 1A-pray=JUNCT=1A=BEN=DECL

- b. Transcribed by Edwin Ko with the assistance of Felice Big Day on 6 November, 2022:

helak um b-akúpee=sh bah-kuxsá=k bas-bilaxpáakee=sh kon
 so.then FILL 1POSS-sibling=DECL 1A-help=SS 1POSS-people=DEF PRO
 mah-kuxssá=k uh ammaa=ítchee=sh xaxúa **kon**
 1A-help=SS FILL REL=good=DEF all **PRO.AGT**
 waaátchim-ma=htaa=(a)k bah-chiwakáa=(a)=wa=ku=k
 be.fortunate-1A=INDIR.CAUS=SS 1A-pray=JUNCT=1A=BEN=DECL

‘And I help my siblings, I help my relatives, and I pray for them so that all good things are bestowed upon them.’

(Theresa Sends Part Home; Theresa_080218_004.wav: 3:31–3:41)

In Example (5a), Felice initially interpreted the bolded form as the ambiguous non-agentive pronominal or focus marker *ko* (see Graczyk 2007: 76–78, 222–223). However, when I revisited the same text with Felice five days later, she reinterpreted this same word as the agentive pronominal *kon*. As a focus marker, *ko* typically follows a noun phrase. As pronominals, however, *kon* and *ko* exhibit anaphoric and endophoric functions, and their position in the clause is based on discourse-pragmatic factors. Thus, in (5a), *ko* could either serve to focus the noun phrase *ammaaítcheesh xaxúa* ‘all good things’ or refer back to the direct object *basbilaxpáakeesh* ‘my family’ introduced in the preceding clause. In (5b), the role of the pronominal *kon* is less clear, since it typically references a third-person agent. Yet the agent of the entire clause chain is first-person. These inconsistencies underscore how listeners can hear different things in different contexts.

Based on Apsáalooke phonology, it is plausible that the form Theresa articulated was actually *ko*, since the following word, underlyingly *baaátchimmaatak*, surfaces here as *waaátchimmaatak*, which is consistent with postvocalic word-initial lenition. Had the form been *kon*, the expected form would be *maaátchimmaatak*, due to nasal assimilation. Nevertheless, *kon* appears before *mahkuxssák* in both interpretations, ostensibly referring to the first-person subject.¹¹ In this context, it can be challenging for academic linguists to suppress the proclivity to view language as a strictly rule-based system. However, it is important to recognize that Apsáalooke discourse remains relatively understudied, and existing analyses have largely relied on written texts, such as materials developed by the Bilingual Materials Development Center and translations of the New Testament, rather than transcripts of spoken discourse.

11 Randolph Graczyk (personal communication, 2025-04-30) also speculates that due to language shift, “the opposition between *kon* ‘subject’ and *ko* ‘elsewhere’ is breaking down in ‘modern’ Crow”. In other words, Graczyk suggests that since the Apsáalooke language is at an advanced stage of language shift, grammatical distinctions are becoming neutralized in the contemporary language (Campbell and Muntzel 1989).

As a result, current studies on Apsáalooke discourse and the analysis presented here may eventually be proven inaccurate, highlighting the need for further investigation into the functions of *ko* and *kon*, which may offer additional insights into Apsáalooke discursive strategies.

4.3 Different community collaborators

To investigate the nature and extent of varying interpretations provided by different community collaborators, we compare two interpretations of the same passage from Theresa's narrative. Example (6a) shows Felice's interpretation of this section, while Example (6b) shows an interpretation from another community collaborator, Jack Real Bird, a Crow man from Lodge Grass, a town on the reservation, born in the early 1970s.

- (6) a. Transcribed by Edwin Ko with the assistance of Felice Big Day on 4 November, 2022:

héeht baakóon baammáshta=k chiín uh baa-wattáa=(a)k
 CONJ peacefully 1A.do.beadwork=SS also FILL 1A-give.away=SS
 ba-láake **láash-dia=wa=htaa=(a)k** um bis-biakelishte
 1POSS-child **name-do=1A=INDIR.CAUS=SS** FILL 1POSS-sister.in.law

baa=wa-kuú=k

INDEF.OBJ=1A-give=DECL

'And I bead, I often give things away, I have every one of my children named, and I give things to my sisters-in-law.'

- b. Transcribed by Edwin Ko with the assistance of Jack Real Bird on 7 November, 2022:

héeht baakóon baammáshta=k chiín uh baa-wattáa=(a)k
 CONJ peacefully 1A.do.beadwork=SS also FILL 1A-give.away=SS
 ba-láake **láash-dia=wa=hkaa=(a)k** um bis-biakelishte
 1POSS-child **name-do=1A=INDIR.CAUS=SS** FILL 1POSS-sister.in.law

baa=wa-kaá=(a)k um bah-chiwakáa=(a)k Ashkísshiluu=o

INDEF.OBJ=1A-give.away=SS FILL 1A-pray=SS Sun.Dance=PL

ba-láa=(a)k awúush-bimmala=k íipche ii=wah-chiwakáa=(a)k ...
 1A-go=SS sweat.lodge-1A.enter=SS pipe INSTR-1A-pray=SS

'And I bead, I often give things away, I have every one of my children named, and I give things away to my sisters-in-law, I pray, I go into the Sun Dance, I enter the sweat lodge, I pray with the pipe....'

(Theresa Sends Part Home; Theresa_080218_004: 2:57–3:15)

The first difference involves the forms *láashdiawahtaak* and *láashdiawahkaak*, provided by Felice and Jack, respectively. Both transcripts contain the indirect causative morpheme, but they surface as distinct forms: in (6a), the indirect causative appears as =*htaa*, but in (6b), it occurs as =*hkaa*.¹² The

12 The Apsáalooke language distinguishes between two types of causatives: the direct *-ee* and indirect causative =*hche*. In the grammar of Apsáalooke, Graczyk (2007: §13.9.1) reports that the indirect causative is =*hche* and the segment *hch*, the preaspirated voiceless postalveolar affricate [htʃ], changes to *hk*, a preaspirated voiceless velar stop [hk], when followed by a non-front vowel. However, it should be noted that he primarily records the most commonly used forms.

hk ~ *ht* alternation is a documented point of variation (see Graczyk 2007: 28, fn.10), and we often find community collaborators interpreting *k* for *t* and vice versa, depending on the variety the individual speaks. A reviewer also commented that they learned from the late Hu Matthews, a linguist and Bible translator who worked with speakers from Lodge Grass, Crow Agency, and Pryor, that “the variant *htaa* was found in Lodge Grass, with *hkaa* elsewhere”. However, in our data, Jack, who was raised in Lodge Grass, produced *hkaa*, while Felice, who is from Pryor, used *htaa*. The reason for this discrepancy is unclear, but we surmise that the variation is more complex now than it was when Matthews, who conducted his work on the Crow language in 1972, proposed the regional dichotomy.

Additionally, the word-final segment *k* can be analyzed as a declarative or same-subject switch-reference marker. We can disambiguate this by examining what has commonly been referred to in the Siouan literature as “ablaut”, a morphophonological process by which certain morphemes induce changes to the final vowel of the verb stem. The same-subject marker is one such morpheme that triggers ablaut, unlike the declarative. In the form given in (6b), *e* undergoes ablaut to *a* (due to the presence of the same-subject marker). This vowel change, in turn, causes *hch* to surface as *ht* or *hk*. If the declarative were present instead, we would expect the form *láashdiawahchek*, without ablaut.

The second difference involves the forms *baawakuúk* and *baawakaák*, contributed by Felice and Jack, respectively. The former ends in the declarative =*k*, whereas the latter contains the same-subject switch-reference =*ak*, generally used to chain clauses featuring the same subject. Again, ablaut aids in distinguishing which marker appears, as in the form *baawakuúk*, which is composed of the verb root *kuú* ‘give’. Since it neither undergoes ablaut nor ends with stem-final -*aa*, we can easily distinguish the ending as a declarative—if it were the same-subject marker, we would expect *baawakuuák* (cf. Graczyk 2007: 174, Ex.24). The declarative thus signals the end of a clause chain. In the form that Jack provided, *baawakaák*, the verb root *keé* ‘give away’ undergoes ablaut to *kaá*. In this case, we can be confident that the final *k* is the same-subject marker and not the declarative.

The same-subject marker indicates that the following clause shares the same subject as the previous one, while a declarative indicates the end of a clause chain consisting of a chronological sequence of events and states. Therefore, in Jack’s interpretation, the passage contains eleven clauses, nine of which are marked with the same-subject marker. In contrast, Felice’s interpretation consists of four clauses, of which three end in the same-subject =*ak* and only the final clause ends with the declarative =*k*.

4.4 Returning to the original speaker

The final case study focuses on a narrative by Charles Yarlott Jr., an Apsáalooke man born in 1980, which we recorded at the Hardin public library. The narrative describes his memories of the high school in Hardin which he attended, and the negative attitudes of *baaschúile* ‘white people’ (a contraction of *baaishdashúile*, lit. those with yellow eyes) toward the Apsáalooke people.¹³

Example (7) contains interpretations by Felice and Charles. Note that approximately six years passed between Charles’s initial account of the history of Hardin and our interpretation session with him.

13 There are several folk etymologies regarding how *baaishdashúile* came to refer to white people. One proposal is that the first European settlers that the Apsáalooke encountered had hazel eyes. Another claims that the White colonizers were afflicted with a condition (e.g., jaundice) that turned the whites of their eyes yellow.

Although both versions draw from the same narrative, they exhibit several differences, which are bolded below.¹⁴

- (7) a. Transcribed by Edwin Ko with the assistance of Felice Big Day on 19 October, 2022:

town kooté hem baaschiile kuhtáa=taahili=**ak** ko
 town like.that FILL white.person right.there=AUG=**SS** PRO.NONAGT
 ahúu=i=k
 many=HAB=DECL
 ‘There used to be many white people in the town.’

- b. Transcribed by Edwin Ko with the assistance of Charles Yarlott on 24 June, 2024:

(hin)né town kooté hem baaschiile kuhtáa=taahilia ko
this town like.that FILL white.person right.there=AUG PRO.NONAGT/FOC
 ahúu=i=k
 many=HAB=DECL
 ‘There used to be many white people in this town.’
 (Charles Yarlott, Jr.; Charles_072418_004.wav: 5:32–5:47)

The first difference is that Felice understood *(hin)né* ‘this; here’ as being part of the previous sentence. In the recording, there was a long 1.2-second pause between *hinné* (which is often reduced to *né*) and the following code-switched item *town*, but there is no perceptible pause between the preceding word and *(hin)né*. Felice thus seems to have taken the word *(hin)né* as an afterthought, resulting in the previous sentence exhibiting right dislocation of *(hin)né*. Its absence from her interpretation in (7a) does not substantially alter the meaning, especially since the narrative was recorded in Hardin, the town in question, so the demonstrative ‘this’ can be inferred from context.

Felice also deviated from Charles in interpreting the adverbial *kuhtáataahilia* ‘right there’ as the predicate *kuhtáataahiliak*. This interpretation most likely contains the same-subject marker rather than the declarative. If it were the declarative, the word would instead surface as *kuhtáataahilik*. In this case, we suspect the following word-initial segment *k* of the following pronominal *ko* influenced Felice’s understanding of the preceding word. We might describe Felice’s interpretation in the following way. Felice believed the speaker, Charles, to have anticipated the following word-initial segment to be identical (i.e., the *k* of *ko*). She assumed that Charles extended the timing of the closure of the word-final segment (i.e., the *k* of *kuhtáataahiliak*), releasing it only when articulating the beginning of the following word. These two differences do not affect the meaning in any significant way. Nevertheless, it is worth noting that interpretations varied between the community collaborator who provided the story and a collaborator encountering it for the first time.

Example (8) is another set of transcripts that exhibit two additional differences, one of which is associated with (dialectal) variation among speakers of Apsáalooke. As noted above, Felice is from Pryor, but Charles was raised in St. Xavier, Montana. These two towns are about a 45-minute drive from each

¹⁴ Graczyk (2007: 366–367) glosses *kuhtée* as ‘specific location,’ analyzing it as the discourse anaphor *ku* plus the postpositional marker *htée*, which refers to a specific place. However, this gloss is somewhat misleading, as *kuhtée* can also refer to distinct points in time (e.g. Graczyk 2007: 367, Ex.26).

other. Since Pryor is somewhat isolated from the other towns on the reservation, speakers from outside of Pryor often comment (favorably) that those from Pryor exhibit some innovations and archaisms in their speech.

- (8) a. Transcribed by Edwin Ko with the assistance of Felice Big Day on 19 October, 2022:

hilihtée am=maa=**íaschil**=uua koon ahóo=sh hilaakée
 right.there REL=INDEF.OBJ=**buy**=PL there many=PST right.now
kala=koot=káashi=ssaa=k
now=like.that=AUG=NEG=DECL

- b. Transcribed by Edwin Ko with the assistance of Charles Yarlott Jr. on 24 June, 2024:

hilihtée am=maa=**ítchil**=uua koon ahóo=sh hilaakée **kalakoón**
 right.there REL=INDEF.OBJ=**buy**=PL there many=PST right.now **then**
 koot=káashi=ssaa=k
 like.that=AUG=NEG=DECL

‘There used to be many stores (lit. where they buy things) right here [in Hardin], but now it’s no longer like that.’

(Charles Yarlott Jr.; Charles_072418_004.wav: 5:37–5:41)

In (8), we observe two main differences between the transcripts. First, the word for ‘store’ is repeated as *ammaaíaschiluua* by Felice in (8a) but as *ammaaítchiluua* by Charles in (8b). Inspecting the recording using Praat, a software for acoustic analysis (Boersma 2001), indicates that the segment of interest is more characteristic of *tch* [tʃ] than *sch* [ʃtʃ]. The Crow dictionary includes the entry for *ammaaíaschiluua* but not *ammaaítchiluua*, raising the question of whether these forms are perfect synonyms. When we followed up with Charles the following day, he offered the sentence in (9), which contains the previously undocumented variants *ammaítchiluua* ‘store’ and *awítchilik* ‘I bought it’. When asked if there were any differences in meaning between the two forms *awíaschilik* and *awítchilik* ‘I bought it’, Charles replied, “it’s the same thing” (personal communication, 2024-06-25).¹⁵

- (9) ammaaítchiluua koon aw-ítchili=k
 store there 1A-buy=DECL
 ‘I bought it at the store.’
 (Charles Yarlott Jr., 25 June, 2024)

The second difference in the transcript presented in (8) involves *kala* = ‘now’ and *kalakoón* ‘then’. In our experience, the former occurs less frequently than the latter in spoken discourse. As Charles’s speech had a rapid succession of *k*-initial syllables, Felice interpreted it as *kala* = instead of Charles’s proposed interpretation *kalakoón*. In our opinion, it is difficult to adjudicate whose interpretation is ‘correct’; rather, each individual’s interpretation merely reflects how they heard and understood it in their own terms.

¹⁵ There are at least two possible ways to analyze these variants. First, the undocumented forms may reflect Charles’s idiolect, which may have been impacted by change during language shift. Second, they may indicate some kind of variation within the community. Further investigation is needed to determine whether this is dialectal or carries indexical links to another social variable.

5 DISCUSSION

Given the variability in interpretation discussed above, whose transcripts should be considered the most reliable or accurate? Rather than privileging one interpretation over another, we advocate for the preservation and archiving of all versions of transcripts. Each version reflects a distinct, yet equally valid, understanding of a given text and may contain insights that would otherwise be lost. By “all versions of transcripts”, we intend the initial draft the researcher produced, any interpretations that linguists may find surprising and perplexing, and even transcripts in which a speaker’s interpretation diverges from the recorded audio.

We further recommend consulting with the original speaker—ideally sooner rather than later.¹⁶ This suggestion contradicts Jung and Himmelmann’s (2011: 206) assertion that “[i]t is not necessary, and in fact often not desirable, to work on transcription with the speaker who appears in the recording”. Their reasons include that these speakers may amend their recorded utterances and may not “[restate] precisely what was actually said”. However, these issues hold true for any speaker, and Jung and Himmelmann (2011: 207–218) even provide examples of interpreters (intentionally) deviating from the original recording (see also Chelliah 2001; Dobrin 2021; Crowley 2007; Haviland 1996; Urban 1996).

There are compelling reasons to involve the original speaker in the interpretation of their own recording. Chief among them is the preservation of the speaker’s voice—defined by Webster (2013: 136) as the “ability to express something in a satisfying manner using all of one’s expressive options”. When the researcher transcribes only the editorializations of other speakers, they risk obscuring or erasing aspects of the recorded speaker’s voice. For instance, in the examples featuring Charles’s recorded narratives, Felice rendered *ammaaítchiluua* ‘store’ as *ammaaíschiluua*. Had we only recorded Felice’s interpretation, we would have inadvertently marginalized Charles’s voice. Although Felice’s interpretation provides valuable insights into inter-speaker variation, it is only through re-engaging with Charles that we were able to maintain Charles’s agency to express himself.

More broadly, these examples highlight the contribution of audience design to the shape of the narratives we recorded. As Bell (1984) explains, audience design defines how speakers organize their speech style to reflect the audience that they assume. In this case, the audience may include both physically present addressees and imagined future listeners. Speech is never produced in isolation—as Dobrin (2012: 22) writes, “[s]peakers are never just talking; they are always also representing themselves to someone”. Schwartz (2021) similarly observes that speakers being recorded are often aware that they serve as representatives of their speech community, embodying not only their own identity but also their culture.

Although misunderstandings can be resolved by asking speakers for additional clarification, this is not always feasible when playing back a recording for a community collaborator to interpret. The speaker whose voice is featured in the recording may not be able to be physically present—they may be unavailable or may have since passed on. Even if the person is available, if significant time has passed since the creation of the initial recording, the speaker may no longer recall their intended meaning. Additionally, speakers might wish to modify what they originally said to better represent themselves at that particular moment in time. These and other factors compound to make the interpretation process prone to some of the types of variation we observed in our transcripts.

16 While we recommend producing a transcript soon after the recording session, we recognize that this is not always feasible—logistical challenges inevitably and unexpectedly arise when conducting fieldwork.

Today, only a few people are literate in Apsáalooke. We do not have the luxury of obtaining written interpretations from a wide range of speakers, nor do we have the resources to train speakers and active listeners (also commonly referred to as passive speakers) to transcribe in an orthography, as is more common in endangered language contexts in Amazonia (Woodbury and England 2004) and Indonesia (Florey and Himmelmann 2009). Even if we can train community members, they may inadvertently alter the content of the recordings to better reflect their vision of how their language and culture should be represented, without much regard for how the original speaker(s) wished to be depicted. This is an occurrence we have commonly observed when working closely with speakers, as they sometimes express a preference to phrase things differently or comment on the ‘incorrectness’ of other speakers’ utterances (see also Dobrin 2021; Chelliah 2001; Haviland 1996; Jung & Himmelmann 2011; Urban 1996; Weber 2021).

Furthermore, working with or training speakers to transcribe and interpret recorded narratives requires resources that could be allocated elsewhere. If a linguist is sufficiently familiar with the language and works independently to transcribe and translate the narratives, that individual might be able to avoid the obstacles arising from conflicting interpretations. However, they may still inadvertently erase and marginalize the original speaker’s voice, since even the linguist’s biases can influence the interpretation process. Even minor decisions, such as how to identify sentence boundaries and whether to include false starts, can alter how the speaker’s voice is conveyed in the transcript. In language revitalization contexts, such transcripts could be viewed as authoritative even though they were created without input from the original speaker and—assuming the linguist is not fluent in the language they are documenting— other speakers of the language.

In the early days of transcribing spoken discourse, before the advent of audio recording technology, transcripts were recorded entirely through dictation, which necessarily precluded the transcription of fast-speech phenomena. Today, written texts from that period are held in high regard, as they are often believed to contain more ‘authentic’ language, especially for languages without known fluent speakers. Some language reclamation practitioners design language classes around such texts (Ko, n.d.). In one example, texts transcribed and written down by James O. Dorsey (1848–1895) in Kanza, a Siouan language, are afforded the same authority and credibility as the speech of a contemporary fluent speaker (for a discussion on the issues surrounding Dorsey’s Kanza materials, see McBride 2009, 2020). When language resources are extremely limited, it behooves one to use whatever materials can be found. Still, caution must be taken when working with past documentation. It is best to approach it with a critical lens.

In each of the case studies presented above, we made implicit and explicit judgments on the various interpretations we received. A reviewer wonders “what transcripts are to be documenting (roughly, production vs. reception)” and “if any criteria are used regarding when to take a stance about which variant to use vs. when to keep multiple variants in parallel”. Ultimately, we believe it depends on one’s research questions and goals. Although we suggest maintaining and archiving all transcripts, we presented only a single speaker’s interpretation on our storytelling mapping platform. We selected each interpretation based on how closely it aligned with the voice of the recorded speaker. For example, we selected the interpretation obtained in person rather than the one created during remote communication over Zoom. However, in this paper, we presented multiple variants in parallel since our goal was to foreground the various interpretations produced by even a single individual.

Of course, not all researchers will want or be able to focus on the quality of the transcripts or their faithfulness to the recordings as we have here. A scholar analyzing naturalistic recordings of conversations

or child-caregiver interactions has the additional challenge of transcribing speech amidst background noise, which often results in unintelligible utterances and difficulty in identifying the turns of each speaker. As Amalia Skilton (personal communication, 2024-07-29) notes, “there will often be multiple plausible hypotheses about what is being said or should have been said (for a child)”. Depending on one’s research questions and methods, it is important to consider whether the quantity or quality of the transcripts should be prioritized, as well as to weigh the potential tradeoffs of each approach.

Just like all transcripts, our transcripts are idealizations of spoken discourse. The audio format inevitably excludes all context, including gestures, space, and posture. Video recording captures only slightly more extralinguistic information. Creating texts out of recorded speech inherently results in a degree of decontextualization (Bauman & Briggs 1990). As individuals interpret the recordings, recontextualization also occurs, leading to the (re)creation of ostensibly limitless interpretations. In this sense, the interpretative act will differ from person to person. These points build on other scholars’ calls for documentary linguists to view their work in an interpretivist paradigm, rather than a positivist one where replication and reproducibility are the goals (e.g., Dobrin 2021; Dobrin and Schwartz 2021; Seidel 2016).

Recent efforts have focused on the development of speech recognition (e.g., Prud’hommeaux et al. 2021) and morphological processing technologies (e.g., Eskander et al. 2019) for endangered (or low-resources) languages to overcome the so-called ‘transcription bottleneck’ and expedite the transcription process. If we already obtain a variety of interpretations from different speakers, how can we expect machines, even with input from linguists, to reach a similarly fine-grained level of performance? Even NLP (or natural language processing) tools for high-resource languages, such as English, have difficulty dealing with variation (for an overview of the potentially negative impacts of language technologies, see D’Arcy & Bender 2023). The issue of variation is further exacerbated in endangered language contexts because speakers in a given community may have very different language acquisition histories.

At the start of this paper, we presented two types of transcripts—a naturalized transcript and a denaturalized transcript. Our decision to format the transcripts in these ways was based on our goals: to serve both the community’s language revitalization and reclamation efforts and the work of any scholars interested in researching the language. At the same time, our approach has consequences that may contravene the goals and aspirations of some members of the community. Significant among these is our use of the standard Apsáalooke orthography, which often fails to denote variation. We also made a conscious decision to include only some paralinguistic features in our transcripts, leaving out expressive features such as pausing and prosody. Our transcripts therefore do not fully showcase the speaker’s voice and instead risk reinforcing a Western bias that privileges referential content over expressive or affective dimensions of speech (Webster 2013; see also Kroskrity 2013). However, by presenting more than one transcript of the same stretch of speech, we aim to contribute to what McDonnell (2018) refers to as *facilitative documentation*: work that is not driven by a specific research question but is designed to support a range of future inquiries, for both the researcher and the community.

The iterative and evolving nature of transcription presents many opportunities for academic and community researchers alike, but only if evidence of this evolution is preserved. What makes language documentation distinct from other disciplines that generate transcripts is the integral role of the archive (Henke & Berez-Kroeker 2016). Based on our experience conducting archival research, it is common to find correspondence and drafts of (un)published works in the collections of earlier scholars. However, there is a very real danger of losing these documents to natural disasters. Today, digital archiving

presents other challenges. Unless printed out, internet correspondence, specifically email, is rarely, if ever, deposited, even though born-digital transcripts—much like their handwritten counterparts—may themselves build on previous sets of transcription (e.g., Debenport 2015; Evans and Sasse 2007). In the digital space, transcripts can be overwritten with the click of a button, leaving no trace of these earlier versions. Thus, we encourage researchers to keep track of changes they may make to transcripts, to prevent the loss of that information.

We also recommend that language archivists adopt incremental archiving, a process by which digital files are uploaded in stages, and subsets are accessioned as they become ready. This model is currently implemented by the California Language Archive (CLA; Garrett et al. 2019). In contrast, other archives require that all files in a deposit, along with complete metadata, be submitted in a single submission. While most, if not all, archives allow modifications to existing metadata, not all allow revision to the content of already accessioned material. For example, the CLA does not allow amendments to accessioned content, including adding to, amending, or removing files from a file bundle, thereby ensuring that previous work is not overwritten.¹⁷

6 CONCLUSION

The first aim of this paper was to discuss and reflect on the challenges of our transcription process. We provided two transcripts, formatted in a naturalized and denaturalized style; the former is more reminiscent of standardized, written discourse, whereas the latter preserves more elements found in spoken discourse, including processes associated with connected speech, hesitation particles, and false starts. The second aim was to show how a single recorded narrative can give rise to a multiplicity of interpretations due both to the means of communication and the subjective nature of the interpretive act itself. We also identified multiple points of variation that scholars of Apsáalooke may wish to explore in the future. We proposed that all drafts of transcripts should be maintained and preserved, including the researcher's initial drafts and those produced in collaboration with speakers that may not exactly align with the recording.

In fact, many of the narratives about personal experiences were documented without explicit connection to any specific research question. Thus, these texts have the potential to both benefit the speech community and facilitate the research of future scholars (McDonnell 2018). For instance, most are considered the standard genre used in sociolinguistic research on variation as they contain extended stretches of speech recounting personal experiences in both Apsáalooke and English (see Sneller & Barnhardt 2023). Despite the fact that significantly more endangered language data has been recorded, research on sociolinguistic variation remains understudied (Seifart et al. 2018; see also Hildebrandt et al. 2017); these texts could enable future studies into this and other topics.¹⁸ That said, we conducted interviews without asking

17 We thank Zachary O'Hagan (personal communication, 2025-07-23) for his insights into the California Language Archive and the archiving practices of other language archives.

18 When I mentioned that some younger speakers of Apsáalooke are reported to speak “like an Elder”, a colleague, Julianne Kapner, remarked that, if this were the case, it would differ from the age-graded variation that is often found for large, urban communities (Labov 1994, 2001; Tagliamonte & D'Arcy 2009). In Apsáalooke society, which is relatively smaller and more rural, grandparents traditionally take a primary role in language socialization. In fact, scholars have shown that categories of age are not fixed but fluid, shaped by the specific community's ideologies and socialization practices (see, e.g., Berman 2014; Meek 2007, 2012; Purkharthofer 2020; Suslak 2009).

follow-up questions, which may preclude the effects of speaker accommodation and rule out opportunities for backchanneling in the language. As Skilton (2017: 111) remarks, recording this back-and-forth could have provided valuable demographic information about the participant, information that I have not been documenting as thoroughly and would be necessary for any variationist study.¹⁹

As a field, we need to more explicitly acknowledge the kinds of variation we encounter and document in our work—variation that, we believe, is not unique to our own Apsáalooke project but likely observed by other language documentarians. By detailing our transcription process, the challenges we faced, and the strategies we employed to address them, we hope to contribute to broader conversations about the transcription of spoken discourse in language documentation. To conclude, we offer a set of considerations that field linguists may find useful when developing transcripts: First, consider your readership and your research goals. As we noted at the outset of this paper, our transcription process was shaped by two primary goals, and we designed our transcripts accordingly. Second, think about providing multiple transcript formats. In our project, we provided two versions to accommodate the needs of different users. Third, be attentive to whose voice—and whose ear—is being privileged in your transcripts. Conflicting interpretations from different individuals are to be expected; when they arise, consider preserving all of them and clearly attributing each interpretation to the individual who provided it. Fourth, keep in mind that standard orthography, while useful for accessibility, can erase socially meaningful variation. Efforts to standardize should be weighed against the risk of introducing negative indexical associations to non-standard forms. Finally, reflect critically on the biases you bring to the transcription process. All researchers carry assumptions and predispositions, and it is essential to recognize and account for them in your work. By taking these considerations into account, we hope that researchers in language documentation will be better equipped to document and represent interpretive variation in future work.

ACKNOWLEDGEMENTS

We are deeply grateful to the Apsáalooke speakers—especially Felice Big Day, Jack Real Bird, Theresa Sends Part Home, and Charles Yarlott Jr.—for generously sharing their personal stories, time, knowledge, and culture with us. We also thank Claire Bown, Randy Graczyk, Julianne Kapner, Zach O'Hagan, Amalia Skilton, and the participants of the 45th Siouan and Caddoan Languages Conference and the Tuesday Seminar Series at the University of Hawai'i at Manoa's Department of Linguistics for their valuable feedback on various stages of this project. This research was supported by the Oswalt Endangered Language Grant and the National Science Foundation Doctoral Dissertation Research Improvement Grant. Finally, we are grateful for the thoughtful comments provided by three anonymous reviewers and extend special thanks to Lise Dobrin, editor of *Language Documentation and Description*, for her insightful guidance and continued encouragement throughout the project. All remaining errors are ours.

COMPETING INTERESTS

The authors have no competing interests to declare.

19 In June 2025, I conducted fieldwork on the Crow Indian Reservation and met with nine of the individuals I had interviewed since 2018. During these visits, I collected information on factors such as age, gender, education, places lived, and upbringing. The list of questions used to elicit potential factors contributing to variation within the community was developed in collaboration with Jack Real Bird.

AUTHOR AFFILIATIONS

Edwin Ko  orcid.org/0000-0002-8275-9859

Linguistics, Yale University, US

Jem Burch  orcid.org/0009-0005-9387-0885

Yale University, US

REFERENCES

- Alden, Arthur, Felice Big Day, Morton Big Medicine, Eugene Deputy, Calvin Jefferson III, Cyle Old Elk, Alma Real Bird, Birdie Real Bird, Jack Real Bird, Theresa Sends Part Home, Riley Singer, Lewis Walks Over Ice, Charles Yarlott Jr., Roberta Yarlott & Edwin Ko. n.d. Crow (Apsáalooke) Field Materials, 2018-17. Survey of California and Other Indian Languages, University of California, Berkeley. <https://doi.org/10.7297/X24Q7S5F>
- Austin, Peter K. 2017. Language documentation and legacy text materials. *Asian and African Languages and Linguistics* 11: 23–44.
- Bauman, Richard & Charles Briggs. 1990. Poetics and performances as critical perspectives on language and social life. *Annual Review of Anthropology* 19: 59–88.
- Begay, Kayla, Justin Spence & Cheryl Tuttle. 2021. Teaching Wailaki: Archives, interpretation, and collaboration. In Adrianna Link, Abigail Shelton & Patrick Sperio (eds.), *Indigenous languages and the promise of archives*, 399–424. Lincoln: University of Nebraska Press.
- Bell, Allan. 1984. Language style as audience design. *Language in Society* 13(2): 145–204.
- Bell, Allan. 2011. Re-constructing Babel: Discourse analysis, hermeneutics and the Interpretive Arc. *Discourse Studies* 13(5): 519–568.
- Berman, Elise. 2014. Negotiating age: Direct speech and the sociolinguistic production of childhood in the Marshall Islands. *Journal of Linguistic Anthropology* 24(2): 109–132.
- Boas, Franz. 1889. On alternating sounds. *American Anthropologist* 2(1): 47–54.
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9–10): 341–345.
- Boland, Julie, Pedro Fonseca, Ilana Mermelstein & Myles Williamson. 2022. Zoom disrupts the rhythm of conversation. *Journal of Experimental Psychology: General* 151(6): 1272–1282.
- Bowern, Claire. 2015. *Linguistic fieldwork: A practical guide*. New York: Palgrave Macmillan.
- Briggs, Charles L. 1986. *Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research*. Cambridge: Cambridge University Press.
- Bucholtz, Mary. 2000. The politics of transcription. *Journal of Pragmatics* 32(10): 1439–1465.
- Cahill, Michael & Keren Rice. 2014. *Developing orthographies for unwritten languages*. Dallas: SIL International.
- Campbell, Lyle & Martha C. Muntzel. 1989. The structural consequences of language death. In Nancy Dorian (ed.), *Investigating obsolescence: Studies in language contraction and death*, 181–196. Cambridge: Cambridge University Press.
- Chelliah, Shobhana L. 2001. The role of text collection and elicitation in linguistic fieldwork. In Paul Newman & Martha Ratliff (eds.), *Linguistic Fieldwork*, 152–165. Cambridge: Cambridge University Press.
- Chelliah, Shobhana L. & Willem J. De Reuse. 2011. *Handbook of descriptive linguistic fieldwork*. Dordrecht: Springer.
- Crowley, Terry. 2007. *Field linguistics: A beginner's guide*. Oxford: Oxford University Press.

- D'Arcy, Alexandra & Emily Bender. 2023. Ethics in linguistics. *Annual Review of Linguistics* 9(1): 49–69.
- De Fina, Anna & Sabina Perrino. 2011. Introduction: Interviews vs. 'natural' contexts: A false dilemma. *Language in Society* 40: 1–11.
- Debenport, Erin. 2015. *Fixing the books: Secrecy, literacy, and perfectibility in indigenous New Mexico*. Santa Fe: SAR Press.
- Dobrin, Lise M. 2012. Ethnopoetic analysis as a resource for endangered-language linguistics: The social production of an Arapesh text. *Anthropological Linguistics* 54(1): 1–32.
- Dobrin, Lise M. 2021. The Arapesh "suitcase miracle": The interpretive value of reproducible research. *Language Documentation and Description* 21: 37–69. <https://doi.org/10.25894/ldd16>
- Dobrin, Lise M., & Saul Schwartz. 2021. The social lives of linguistic legacy materials. *Language Documentation and Description* 21: 1–36. <https://doi.org/10.25894/ldd12>
- Epps, Patience L., Anthony K. Webster & Anthony C. Woodbury. 2023. Documenting speech play and verbal art: A tutorial. In Peter Jenks and Lev Michael (eds.), *Key topics in language documentation and description* (Language Documentation & Conservation Special Publication 26): 175–241. Honolulu: University of Hawai'i Press. <https://hdl.handle.net/10125/75035>
- Eskander, Ramy, Judith L. Klavans & Smaranda Muresan. 2019. Unsupervised morphological segmentation for low-resource polysynthetic languages. In *Proceedings of the 16th Workshop on Computational Research in Phonetics, Phonology, and Morphology*, 189–195. Florence, Italy: Association for Computational Linguistics.
- Evans, Nicholas & 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. <https://doi.org/10.25894/ldd261>
- Fairclough, Norman. 1989. *Language and power*. London: Longman.
- Florey, Margaret & Nikolaus P. Himmelmann. 2009. New directions in field linguistics: Training strategies for language documentation in Indonesia. In Margaret Florey (ed.), *Endangered languages of Austronesia*, 121–140. Oxford: Oxford University Press.
- Foley, William A. 2002. Field methods. In Kirsten Malmkjær (ed.), *The Linguistics Encyclopedia*, 131–137. London and New York: Routledge.
- Garrett, Andrew, Edwin Ko, Julia Nee, Zachary O'Hagan & Ronald Sprouse. 2019. 'Pre-archiving' with the California Language Archive: Incremental archiving and early ongoing curation. Paper presented at the 6th International Conference on Language Documentation and Conservation (ICLDC). Online at <https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/898a2aef-83c5-4d5a-9744-66580cfdecbb1/content>. Accessed 2025-11-29.
- Goddard, Ives. 1976. Philological approaches to the study of North American Indian languages: Documents and documentation. In Thomas A. Sebeok, *Native languages of the Americas*, 73–91. New York: Springer.
- Goffman, Erving. 1981. *Forms of talk*. Philadelphia: University of Pennsylvania Press.
- Gordon, Ray & Randolph Graczyk. n.d. *A dictionary of Crow*. Unpublished manuscript.
- Graczyk, Randolph. 2007. *A Grammar of Crow*. Lincoln: University of Nebraska Press.
- Haviland, John B. 1996. Text from talk in Tzotzil. In Michael Silverstein & Greg Urban (eds.), *Natural histories of discourse*, 45–78. Chicago: University of Chicago Press.
- Henke, Ryan & Andrea L. Berez-Kroeker. 2016. A brief history of archiving in language documentation, with an annotated bibliography. *Language Documentation & Conservation* 10: 411–457. <http://hdl.handle.net/10125/24714>

- Hepburn, Alexa, and Galina B. Bolden. 2013. The conversation analytic approach to transcription. In Jack Sidnell and Tanya Stivers (eds.), *The Handbook of Conversation Analysis*, 57–76. Chichester: Wiley-Blackwell.
- Hildebrandt, Kristine A., Carmen Jany & Wilson Silva. 2017. Introduction: Documenting variation in endangered languages. In Kristine Hildebrandt, Carmen Jany & Wilson Silva (eds.), *Documenting Variation in Endangered Languages* (Language Documentation & Conservation Special Publication 13): 1–5.
- Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for. In Jost Gippert, Nikolaus Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 1–30. Berlin: Mouton de Gruyter.
- Himmelmann, Nikolaus. 2018. Meeting the transcription challenge. In Bradley McDonnell, Andrea Berez-Kroeker & Gary Holton (eds.), *Reflections on language documentation 20 years after Himmelmann 1998* (Language Documentation & Conservation Special Publication 15): 33–40. Honolulu: University of Hawai'i Press. <http://hdl.handle.net/10125/24806>
- Hymes, Dell. 2004. *Ethnography, linguistics, narrative inequality: Toward an understanding of voice*. London: Taylor & Francis.
- Hymes, Virginia. 1987. Warm Springs Sahaptin narrative analysis. In Joel Sherzer & Anthony C. Woodbury, *Native American discourse: Poetics and rhetoric*, 62–102. Cambridge: Cambridge University Press.
- Innes, Pamela. 2010. Ethical problems in archival research: Beyond accessibility. *Language & Communication* 30(3): 198–203.
- Jung, Dagmar, and Nikolaus P. Himmelmann. 2012. Retelling data: Working on transcription. In Geoffrey Haig, Nicole Nau, Stefan Schnell & Claudia Wegener (eds.), *Documenting endangered languages: Achievements and perspectives*, 201–220. Berlin: Mouton de Gruyter.
- Ko, Edwin. n.d. Incorporating texts into language revitalization and reclamation efforts: Prospects and challenges. Unpublished manuscript.
- Kowal, Sabine & Daniel C. O'Connell. 2014. Transcription as a crucial step of data analysis. In Uwe Flick (ed.), *The SAGE handbook of qualitative data analysis*, 64–78. London: Sage.
- Kroskrity, Paul V. 2013. Discursive discriminations in the representation of Western Mono and Yokuts stories: Confronting narrative inequality and listening to indigenous voices in Central California. *Journal of Folklore Research: An International Journal of Folklore and Ethnomusicology* 50(1–3): 145–174.
- Labov, William. 1994. *Principles of linguistic change. Volume 1: Internal factors*. Oxford: Blackwell.
- Labov, William. 2001. *Principles of linguistic change. Volume 2: Social factors*. Oxford: Blackwell.
- McBride, Justin. 2009. Orthography and ideology: Examining the development of Kaw writing. In Daisy Rosenblum & Carrie Meeker (eds.), *Santa Barbara papers in linguistics: Proceedings from the 12th annual workshop on American Indigenous languages*, 20: 30–45.
- McBride, Justin. 2020. Considerations for the continuation of a Kansa corpus. In Ryan M. Kasak, Samiron Dutta & Justin McBride (eds.), *Proceedings of the 39th Siouan and Caddoan languages conference*: 39–51.
- McDonnell, Bradley. 2018. Reflections on linguistic analysis in documentary linguistics. In Bradley McDonnell, Andrea Berez-Kroeker & Gary Holton (eds.), *Reflections on language documentation 20 years after Himmelmann 1998* (Language Documentation & Conservation Special Publication 15): 191–200. Honolulu: University of Hawai'i Press. <http://hdl.handle.net/10125/24820>

- Meek, Barbra A. 2007. Respecting the language of elders: Ideological shift and linguistic discontinuity in a Northern Athapascan community. *Journal of Linguistic Anthropology* 17(1): 23–43.
- Meek, Barbra A. 2012. *We are our language: An ethnography of language revitalization in a Northern Athabaskan community*. Arizona: University of Arizona Press.
- Mithun, Marianne. 1989. The incipient obsolescence of polysynthesis: Cayuga in Ontario and Oklahoma. In Nancy Dorian (ed.), *Investigating obsolescence: Studies in language contraction and death*, 243–258. Cambridge: Cambridge University Press.
- Mondada, Lorenza. 2007. Commentary: Transcript variations and the indexicality of transcribing practices. *Discourse Studies* 9(6): 809–821.
- Moore, Robert. 2013. Reinventing ethnopoetics. *Journal of Folklore Research: An International Journal of Folklore and Ethnomusicology* 50(1–3): 13–39.
- Nathan, David & Peter K. Austin. 2004. Reconceiving metadata: Language documentation through thick and thin. *Language Documentation and Description* 2: 179–188. <https://doi.org/10.25894/ldd299>
- Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs & Bambi Schieffelin (eds.), *Developmental Pragmatics*, 43–72. New York: Academic Press.
- Purkarthofer, Judith. 2020. Intergenerational challenges: Of handing down languages, passing on practices, and bringing multilingual speakers into being. In Andrea C. Schalley & Susana A. Eisenclas (eds.), *Handbook of home language maintenance and development: Social and affective factors* (Handbooks of Applied Linguistics 18): 130–150. Berlin: De Gruyter Mouton.
- Prud'hommeaux, Emily, Robbie Jimerson, Richard Hatcher & Karin Michelson. 2021. Automatic speech recognition for supporting endangered language documentation. *Language Documentation & Conservation* 15: 491–513. <http://hdl.handle.net/10125/74666>
- Schwartz, Saul. 2021. Legacy materials and cultural facework: Obscenity and bad words in Siouan language documentation. *Language Documentation and Description* 21: 166–198. <https://doi.org/10.25894/ldd21>
- Seidel, Frank. 2016. Documentary linguistics: A language philology of the 21st century. *Language Documentation and Description* 13: 23–63. <https://doi.org/10.25894/ldd154>
- Seifart, Frank, Nicholas Evans, Harald Hammarström & Stephen C. Levinson. 2018. Language documentation twenty-five years on. *Language* 94(4): e324–e345.
- Silverstein, Michael & Greg Urban. 1996. *Natural histories of discourse*. Chicago: University of Chicago Press.
- Skilton, Amalia. 2017. Three speakers, four dialects: Documenting variation in an endangered Amazonian language. In Kristine Hildebrandt, Carmen Jany & Wilson Silva (eds.), *Documenting variation in endangered languages* (Language Documentation & Conservation Special Publication 13), 94–115. Honolulu: University of Hawai'i Press.
- Sneller, Betsy & Adam Barnhardt. 2023. Sociolinguistic prompts in the 21st century: Uniting past approaches and current directions. *Language and Linguistics Compass* 17(3): e12484.
- Suslak, Daniel F. 2009. The sociolinguistic problem of generations. *Language & Communication* 29(3): 199–209.
- Tagliamonte, Sali A. & Alexandra D'arcy. 2009. Peaks beyond phonology: Adolescence, incrementation, and language change. *Language* 85(1): 58–108.

- Urban, Greg. 1996. Entextualization, replication, and power. In Michael Silverstein & Greg Urban (eds.), *Natural histories of discourse*. Chicago: University of Chicago Press.
- Weber, Tobias. 2021. Philology in the folklore archive: Interpreting past documentation of the Kraasna dialect of Estonian. *Language Documentation and Description* 21: 70–100. <https://doi.org/10.25894/ldd18>
- Webster, Anthony K. 2013. The validity of Navajo is in its sounds: On Hymes, Navajo poetry, punning, and the recognition of voice. *Journal of Folklore Research: An International Journal of Folklore and Ethnomusicology* 50(1–3): 117–144.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann & Han Sloetjes. 2006. ELAN: A professional framework for multimodality research. In *5th international conference on language resources and evaluation (LREC 2006)*: 1556–1559.
- Woodbury, Anthony C. 2010. Language documentation. In Peter K. Austin & Julia Sallabank (eds.), *The Cambridge handbook of endangered languages*, 159–186. 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. <https://doi.org/10.25894/ldd161>
- Woodbury, Anthony C. & Nora C. England. 2004. Training speakers of indigenous languages of Latin America at a US university. *Language Documentation and Description* 2: 122–139. <https://doi.org/10.25894/ldd295>