Academic-community collaboration:
Data collection, linguistics research, and language teaching[1]

Mizuki Miyashita
The University of Montana

Rosella Many Bears
Aahsaopi Elementary School

This paper describes our collaborative research project on video-recording spoken Blackfoot language. The goal of this project is multi-dimensional: (i) to video record language use in everyday situations, (ii) to collect data for linguistic analysis, and (iii) to use the data for language teaching. The general structure of the project includes: fieldwork, video-recordings collected, transcription, and interlinear analysis. We believe that collecting video-recordings of spoken Blackfoot capture various ways the language is used that are important but difficult to teach without appropriate context. Since the data capture language in use, it also contributes to increasing knowledge in related linguistics fields such as sociolinguistics and linguistic anthropology.

1 Introduction

In this paper, we will present our collaboration project in Blackfoot language documentation. The goal of the project is twofold: The first is to show a successful example case of collaboration between an academic researcher and a community educator in the area of Blackfoot language documentation. The other is to share with the wider community our project of video-recording uses of the Blackfoot language in real-life situations. The collected data then will be used for linguistic analysis and for applications in Blackfoot teaching. We also hope to promote collaborative projects among linguists and language community members in language documentation and revitalization.

This paper is organized as following. First, we describe our language documentation project which consists of data collection, transcribing, and interlinear analysis. Then we will discuss the significance of our project

[1] We would like to thank Darrell R. Kipp, the director of Piegan Institute. Donald Frantz, Delores Many Bears, Louise Giebel, and Leo & Kristen Kipp. We would also like to thank audience at the ICSNL 44. This project is partially supported by the Native Voice Endowment from the Endangered Language Fund and University of Montana Small Grant. All errors are mine.

regarding applications in the fields of language preservation, linguistics research and language education. Then the conclusion follows.

2 Documentation Project

Our collaborative project's primary goal is to document naturally occurring speech in Blackfoot. We recognize that there are multiple ways to accomplish this kind of documentation. For example, some linguistics choose to write descriptive grammars based on conducting elicitation with native speakers, while some choose to simply record speech such as narratives or conversations. These are both valuable documentation resources, and we believe that the latter type of documentation must include transcription and interlinear analysis, since these require working with native speakers. This concept is illustrated in Figure 1 below:

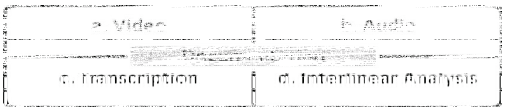


Figure 1

In the following sections, we describe our documentation project which is divided into three stages: data collection, transcription and interlinear analysis.

2.1 Data Collection

We conducted video-recording Blackfoot speech in May 2009. The recording took place at the Piegan Institute in Browning, Montana, United States. The image recording equipment we used was a Canon VIXIA-11. This model is good for a home camera, but the sound quality is not as good as other audio equipment made for sound recording purposes. Therefore, we also used an audio-recorder, Zoom H2. The image and sounds will be integrated.

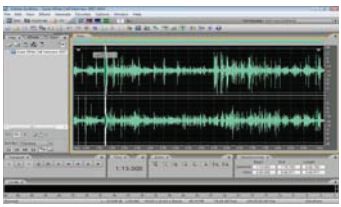
Previously, we had recorded conversations among native speakers while they were sitting at a table, so we are able to document their discussions or gossips. This time, however, we were interested in capturing speech while they are engaging in some actions, and in this particular project, we asked a native speaker to spend some time with a baby. Our recording included the speaker changing the baby's diaper (Figure 2), naturally occurring speech in a normal life activity.

Figure 2

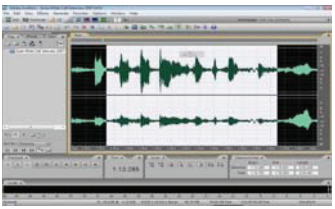
2.2 Transcription

We used the audio file to make transcription. The software used was Adobe Audition 3®, which was designed for music mixing purposes. This program made it easy to transcribe the audio file for the following reasons: The sound recording is stored in a form of digital sound file (.wav), and this can be played by regular audio-playing-software such as Microsoft Media Player® or QuickTime®. As for transcription purpose, it is necessary to play only a very short part of the recording and to repeat the same section multiple times until phrases are recognizable and transcribable. In order to do so, we needed to use a program that enables us to do this. Praat, free software for acoustic phonetics analysis, is capable of doing so, but since the program was primarily designed for phonetic analysis, finding a small section to be analyzed is not very convenient when dealing a large sound file. When using Audition, you can stretch and squeeze the view by using a roller on a mouse. On the other hand, Praat involves more steps (or multiple mouse-clicking) to zoom in selected parts. Moreover, while a mouse must be clicked in Praat to play the selected portion of the file, Audition plays it by pressing the space bar. This is convenient for playing the same part repeatedly. This can also be made into a loop. Furthermore, if a speech is too fast to transcribe, the file can be easily doubled in time to slow down. Thus, Audition lets us find a spot quickly especially when the sound file is large; therefore it saves time for the process of transcribing a long speech.

To give the reader an idea of the style of the user-interface in Audition, Figure 3 shows the screenshots of the sample ten-minute-long sound file played through Adobe Audition 3 (Figure 3 a & b) and Praat (c & d). Both show the screen showing entire 10 minute-long sound wave forms and 2 second-long selected wave forms.

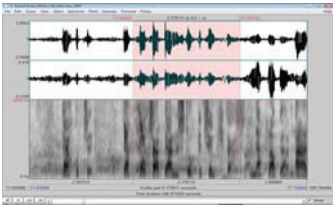


(a) Audition All



(b) Audition Zoomed

(c) Praat All



(d) Praat Zoomed

Figure 3

We transcribed the sound potion of the video-recording by collaborative teamwork Mizuki, a linguist, and the first author, controlled the program to play sound through the computer, while Rosella, the second author who is a native Blackfoot speaker, transcribed and translated the speech by writing down words on a notepad. She has self-taught the official Blackfoot orthography, and it was easier for her to write phrases down by herself than to let a linguist write them down. Based on the notes that Rosella made, Mizuki then typed the transcription/translation, and saved them into a word file. During this process, we made an effort to preserve the exact speech production of the speaker. The example transcription and translation is shown in (1) below.

(1) Transcription and Free Translation

Line	Transcription	Free Translation
1	takisawaistsa naa issitsimaan	I will change this baby's diaper
2	kiakisawaistso ha	I am going to change your diaper. Okay?
3	issitsimaan tsa kitanikkoo	Baby, what is your name?
4	yaa taannatsiistaa	Wow, you had a big poop.

2.3 Interlinear Analysis

The next step of the documentation project was to complete an interlinear analysis of the recorded and transcribed speech. Interlinear analysis is a morphological analysis that finds the smallest meaningful units that comprise a word/phrase, and provide gloss or grammatical function for each part. Interlinear analysis for this project was made based on the available Blackfoot language

materials such as the dictionary (Frantz & Russell 1995), a Blackfoot vocabulary list (Uhlenbeck & Van de Gulik 1930), a grammatical description (Frantz 1991), and English free translations given by Rosella at the time of transcribing, as described in the previous section. This process was challenging because a word often consists of multiple morphemes or affixes, especially since Blackfoot is a polysynthetic language. Also, spoken forms may not match the reference form that is found in the dictionary, and likewise the free translation may not correspond to the literal meaning. For these reasons, the process of interlinear analysis tends to be time-consuming. In addition, Blackfoot undergoes various phonological changes; therefore, one must consider the underlying phonological forms in order to find reference forms of affixes. For example, *tsa kitanikkoo* "what is your name?" includes a verb root *anit* 'say.' But the surface form of this root occurs as *anik* (see Line 3a-b).This is because the last sound [t] of the root *anit* changes to [k] when followed by another [k], following the phonological rule described by Frantz (1966).

(2) below is the transcription of (1) above with the interlinear analysis added below the transcribed lines. The first line (a) is the transcription, the second line (b) is the morpheme-base analysis, and the third line (c) is the gloss of each morpheme. Completing this process requires some linguistics background such as identifying bound morphemes that are not necessarily consciously recognized by native speakers. During this process, we included grammatically important information that is not manifested in the actual speech. For example, the morpheme that shows involvement of a first person participant is *nit*. As shown in Line 1a and Line 1b in (2), the actual speech lacks the initial two phonological segments [ni]. In regular speech, dropping of this [n] and its following [i] is common among native speakers. According to Frantz (p.c.) dropping of person marker is observed only with the first person and affirmative expressions, and never with the second person [k]. This observation has not yet been formally analyzed in the literature. We believe that making a good interlinear analysis of the transcription will allow us to make further linguistic analyses of natural speech in Blackfoot.

(2) Interlinear Analysis

Line		Transcription & Interlinear	Free Translation
1	a.	takisawaistsa naa issitsimaan	I will change this baby's diaper
	b.	nit+áak+isawai'tsi naa issitsimaan	(lit: I will change this baby's state)
	c.	1+ future+ change.vta this baby	
2	a.	kiakisawaistso ha	I am going to change your diaper. Okay?
	b.	kit+áak+isawai'tsi+o ha	
	c.	2+fut+change.vta+dir okay	
3	a.	issitsimaan tsa kitanikkoo	Baby, what is your name?
	b.	issitsimaan tsa kit-ánit+k+oo	
	c.	baby.nan what.und 2-say+inv+?	
4	a.	yaa taannatsiistaa	Wow, you had a big poop.
	b.	yaa taannatsi+i'staa	
	c.	wow ? + defecate	

3 Significance

Our documentation project involved three steps: audio-video recording, transcription of the recordings, and interlinear linear analysis of the transcription. We expect these materials to contribute to the areas of language preservation, linguistics research, and language education. Contribution to the area of language preservation is made by the audio-video files, as well as the transcriptions and the interlinear analysis. Formal Linguistics also benefits from this work: phonetics/phonology benefits from the audio files, morphology, syntax and discourse analysis from the interlinearized transcriptions. Video files also provide data for non-verbal communication analysis. The Field of education also benefits from these materials. Also, data concerning speakers' non-verbal expressions, discourse strategies, and/or cultural norms in communication can be gleaned from the video files. Finally, our transcriptions can help language teachers' literacy development because they can study the mechanics of the language by studying the interlinear analysis.

4 Conclusion

In this paper, we have described our academic-community collaboration project of Blackfoot language documentation. Many endangered language communities wish to revitalize their languages. It is significant for a community to take leadership, since there is a limit in what linguists can do (Grenoble 2009). Linguists can help communities, but successful collaboration occurs when goal-

differences between the two parties are identified and mutually respected (Miyashita and Crow Shoe 2009). It may be impossible to find a universally effective method of collaboration that is equally beneficial to both groups. However, collaboration allows both groups to share the same data for various purposes. As shown in the flowchart below (Figure 4), such collaboration creates opportunities for linguists and native speakers to exchange different skills through their work. Native speakers of course provide data for linguistic analysis, and linguists may introduce new technology to the community. The documentation then becomes part of their language preservation efforts, and will be used for linguistics research and creation of teaching materials. These materials can be disseminated through conference presentations and/or publication. In conclusion, we highly encourage collaborative work between academia and communities for language documentation projects.

The scheme of the collaboration and mutual benefits are illustrated in figure 4. We believe that collaboration between the language community and academia is very important, as has already been discussed as Rice (2009) claims that these two "solitudes" must come to a mutual recognition that linguists cannot do language revitalization by themselves.

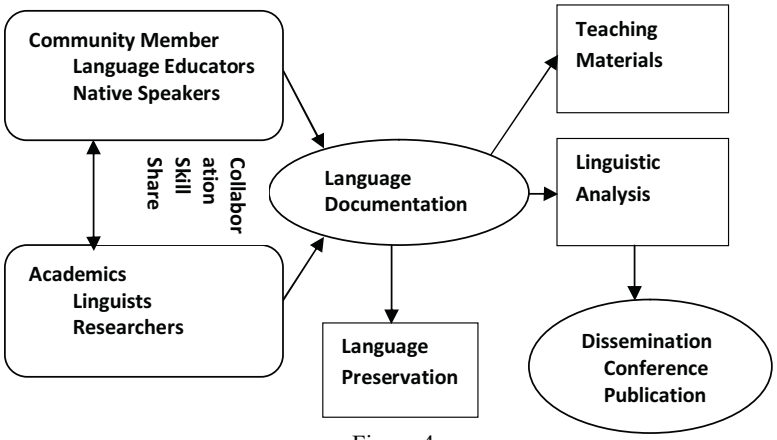


Figure 4

References

Frantz, Donald G. 1966. "Person indexing in Blackfoot." International Journal of American Linguistics 32.50-58.

Grenoble, Lenore. 2009. "Linguistic Cages and the Limits of Linguists." In Indigenous Language Revitalization: Encouragement, Guidance & Lessons Learned, ed. Jon Reyhner & Louise Lockard, pp. 61-69. Flagstaff: Northern Arizona University.

Frantz, Donald G. 1991. Blackfoot Grammar. Toronto: University of Toronto Press.

Frantz, D. and Norma J. Russel. 1989. Blackfoot Dictionary of Stems, Roots, and Affixes. U Toronto.

Rice, Karen. 2009. “Must There Be Two Solitude?” In Indigenous Language Revitalization: Encouragement, Guidance & Lessons Learned, ed. Jon Reyhner & Louise Lockard, pp. 37-59. Flagstaff: Northern Arizona University.

Miyashita, Mizuki, and Shirlee Crow Shoe. 2009. “Blackfoot Lullabies and Language Revitalization.” In Indigenous Language Revitalization: Encouragement, Guidance & Lessons Learned, ed. Jon Reyhner & Louise Lockard, pp. 183-190. Flagstaff: Northern Arizona University.

Rice, Karen. 2009. “Must There Be Two Solitude?” In Indigenous Language Revitalization: Encouragement, Guidance & Lessons Learned, ed. Jon Reyhner & Louise Lockard, pp. 183-190. Flagstaff: Northern Arizona University.

Uhlenbeck, Christian C. & Van Gulik, Robert H. 1930. An English-Blackfoot Vocabulary: Based on Material from the Southern Peigans. Uitgave van de Koninklijke Akademie.