Documenting multimodal interaction: workflows, data management, and archiving

Daisy Rosenblum, University of California, Santa Barbara
Olivia N. Sammons, University of Alberta

“Documenting Conversation”
LSA Annual Meeting: January 2-5, 2014
Acknowledgements

Verna DeMontigny, Mervin Fleury, Norman Fleury, Victoria Genaille, Grace Ledoux-Zoldy, Harvey Pelletier (Michif Elders); David Beck (Univ. of Alberta), Nicole Rosen (Univ. of Manitoba)

Lillian Johnny, Beverly Lagis, Gertrude Robertson, Ernest Scow, Spruce Wamiss, Elizabeth Cadwallader, Mikael Willie, (Kwakwaka’wakw Nation); Marianne Mithun (UCSB), Patricia A. Shaw (UBC)

Funding provided by the Hans Rausing Endangered Languages Project: IGS 0151, 2011 – 2014 (Sammons); IGS 0187, 2012-2014 (Rosenblum)

Further support from the University of California, Santa Barbara, the Jacobs Research Funds, the Phillips Fund for Native American Research, the Oswalt Fund for Fieldwork on Endangered Languages.
Managing conversational data

• Easy to capture, challenging to transcribe
• BOLD: Basic Oral Language Documentation
  (Reiman 2010, Bird 2012)
• Adapting BOLD to transcription of conversation
  – Integrating ELAN
  – Allowing onsite written annotation
  – Attention to archiving
Overview

• Background

• Workflow
  1. Recording
  2. Segmentation
  3. Oral annotation
  4. Transcription
  5. Archiving

• Case studies
  – Michif
  – Kwak’wala

• Discussion & conclusions
Background
“Transcribing narrative and conversational speech is a core activity of all linguistic fieldwork, though one of the less attractive ones. Neither linguists nor speakers are generally very keen to spend long hours on this task”

(Jung and Himmelmann 2011:201).
“As a rule of thumb, never estimate less than a 4:1 ratio between transcription and recordings. That is, for each hour of recording, it will take the linguist at least four hours to transcribe the data, if they are already highly familiar with the language and its sounds. If you are just beginning in a new language with unfamiliar sounds, the ratio is more likely to be 5:1 or 6:1, or even more, until you get used to transcribing this particular language.” (Sakel and Everett 2012:107).

“As a rough estimate, however, you can probably expect to spend something like twenty times the length of the original text on your initial rough transcription with the help of a native-speaker language-helper” (Crowley 2007:138).
Basic Oral Language Documentation

This project is recording and transcribing indigenous languages of Papua New Guinea, using voice recorders donated by Olympus. Papua New Guinea is home to over 800 languages, many with few remaining speakers, and many with minimal linguistic documentation. The work is being done by university staff and students who speak the local languages.

We are collecting narratives, dialogues and songs for 100 languages, using the technique of "Basic Oral Language Documentation" (BOLD). Materials will be freely available for non-commercial use. The project was originally scheduled to run for one year, from 21 February 2010 (UNESCO International Mother Language Day), but was extended by a further year. Activities are continuing in 2012.

In collaboration with staff at the following institutions:
• “a methodology for documenting languages that minimizes the use of high-cost means of recording comments on recorded language data (written annotation), focusing instead on making low-cost means (oral annotation) more effective” (Reiman 2010).

• **BOLD:PNG**

  **Audio Capture**: 10 hours

  **Oral Annotation**: 1 hour respoken and orally translated onto another recorder (10 hours)

  **Transcription**: 0.1 hour of selected recordings transcribed and translated into notebook. (10 hours)
Adapting BOLD methodology

- 3 types of annotation (Reiman: 256)
  - Careful speech
  - Phrasal translation
  - Analytical comments
- Add preparation for transcription session: segment in ELAN
- Allow simultaneous written annotation and analysis during session, in ELAN and field notebook
- Attend to archiving questions related to:
  1. Conversational data: how to handle sensitive material?
  2. Status of oral annotation recordings: treat as primary or secondary data?
ELAN

- Time-aligned transcription software
- Tiers: researcher-defined annotation layers
- Segmentation and playback
- Search
<table>
<thead>
<tr>
<th>No.</th>
<th>Type1: text</th>
<th>Type2: translation</th>
<th>Type3: notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1...</td>
<td>ashkaw kwayesh kiikashkitaan epiihtonamaan anima la ku d-, uh</td>
<td>sometimes you could peel it good, that tail, uh</td>
<td>ashkaw - sometimes; kiikashkitaan - you could; epiihtonamaan -</td>
</tr>
<tr>
<td>1...</td>
<td>la ku di kastor anima</td>
<td>that beaver tail</td>
<td></td>
</tr>
<tr>
<td>1...</td>
<td>ekwa</td>
<td>and</td>
<td></td>
</tr>
<tr>
<td>1...</td>
<td>kipaashin</td>
<td>you would dry it</td>
<td></td>
</tr>
<tr>
<td>1...</td>
<td>apres sa kapa-kapaashtekte ekwa kishiipahin ekoti kekwe</td>
<td>after that, when it was dry, you would comb/brush it there</td>
<td>kapaashteke - when it's dry; VD: kishiikahin; kishiipahin - you brushed/combed it</td>
</tr>
<tr>
<td>1...</td>
<td>lii mitten okik ka-oshahakik, uh</td>
<td>when I made these mittens, uh</td>
<td>okik - these; ka-oshahakik - when I made</td>
</tr>
</tbody>
</table>
Research contexts
Michif - Language situation

• Contact language – Romance (Canadian French) and Algonquian (Plains Cree, Saulteaux)

• Highly endangered - Roughly 600 speakers throughout Canada, less than 1,000 speakers total (Lewis 2009)

• High degree of variation across speakers/communities

• Spoken primarily in Manitoba, Saskatchewan, Alberta, North Dakota, Montana
Documenting Michif Variation
2011 – 2014

Corpus of spoken Michif:
– Observed communicative events: conversations, stories
– Staged communicative events: frog story, pear film, toy game, etc.

(Lüpke 2010)
Kwak’wala


• Active community-based revitalization and restoration efforts

• Polysynthetic, fusional, complex phonology (42 consonants)

• Spoken on Vancouver Island and surrounding mainland British Columbia
Kwak’wala

5 Dialects:
• Kwak’wala
• ‘Nak’wala
• Guc’a
• Tl’atl’asik’walala
• Lik’wala
Building a corpus of interactive speech in Kwak'wala

• Majority of existing Kwak’wala corpus: pre-digital, monologic, narrative (Boas, Hunt)
• How do people talk to each other?
• ~15 hours of conversation recorded; 5 hours translated and transcribed (varying stages of annotation)
• Topics: small talk, embedded narrative, residential school experiences, relocation

Outputs: curriculum materials, linguistic analysis, dissertation
Workflow
recording → data management → segmentation → oral annotation → transcription → archiving
recording → data management → segmentation → oral annotation → transcription → archiving
Recording - Michif

RODE NT4

Canon VIXIA HF S30

Marantz PMD 661

Countryman B3 lavalier microphones

Canon WD-H58 wide converter lens
Sample recording set-up
Recording – Kwak’wala

• Zoom H4N: 4 track setting
• AT Pro-70 lavalier mic
• Canon XA-10
Recording conversations

- Number of speakers
  - Kwak’wala: 2 speakers
  - Michif: 2 – 4 speakers
- Reminder to speakers that recordings will be made public
- Audio metadata
recording → data management → segmentation → oral annotation → transcription → archiving
Data management

- Transfer
- Rename files

- Metadata: preserves links between files
recording → data management →

segmentation → oral annotation →

transcription → archiving
Segmentation

- ELAN segmentation mode: ‘chunk’ recording in preparation for oral annotation session

- Efficient playback
- Quick navigation within primary recording
Segmentation

• What size chunks?

Intonation Units? Clauses? Sentences?
recording → data management → segmentation → oral annotation → transcription → archiving
Oral Annotation

- Transcription session is audio-recorded
- Explain process to annotator/translator
- If sensitive material ask speakers to indicate the beginning and end of sections which they would like to leave untranslated or make inaccessible to archive audiences
- Oral metadata: Filename or ID of source text, Starting time stamp
- Re-speaking and phrasal translation
  1. Play small chunk from source recording in ELAN
  2. Annotator/translator “re-speaks” the chunk (clause, phrase, sentence)
  3. Annotator/translator provides free translation
Set-up

• Laptop
• Good speakers or headphone splitters
• Audio recorder
• Good microphone

Beverly Lagis
Kingcome Inlet, BC
Oral annotation

• How many consultants in a transcription session?

  Option 1: All of speakers from source recording
  Pro: Speakers re-speak and translate their own speech
  Con (?): Reduced efficiency. Expensive!

  Option 2: One speaker
  Pro: Streamlined. Conserves speakers’ time and your funding. Some annotators are especially skillful.
  Con (?): Increased variation. Annotators are translating and re-speaking others’ speech.
Oral annotation

- ‘Respeaking’: not a simple task
- Translating: not a simple task
Different strategies

1. Oral only: no written annotations (BOLD)
2. Free translation only: enter into ELAN
3. Transcription: fieldnotes or ELAN
recording → data management → segmentation → oral annotation → transcription → archiving
Transcription: Post-trip

• Add transcriptions to primary .eaf file
• Where should additional information from transcription sessions be annotated?
• Train and supervise transcription assistants: Local Remote
• Segment oral annotation sessions in ELAN
• SayMore: Data management and processing for oral annotations
recording ➔ data management ➔
segmentation ➔ oral annotation ➔
transcription ➔ archiving
Post-production: metadata and archiving

- Archive .wav, (.mov), .eaf and fieldnote PDF files
- Sensitivity and access:
  - Process marked sections: add noise or scramble
  - Archive two copies unless original is recoverable
- Maintaining links between source and secondary data: Metadata
- Where should recorded annotation sessions be archived?
  - Recording date?
  - With original session?
Michif

- Average conversation session 2 hrs.
- Originally did full written transcription and translation with speaker
  - Roughly 1 min. per hour for multi-speaker recordings (on a good day)
  - Over 60 hours of material to transcribe
    - Approx. 3600 hours to arrive at a rough transcription
    - ~ 90 weeks transcribing full-time (8 hours a day, 5 days a week)
Harvey Pelletier, Mervin Fleury, & Norman Fleury
Brandon, MB
Michif oral annotations

Experimentation with oral annotations:

– Same speakers as in original recording:
  • Not always feasible due to travel, hearing issues, speaker availability, etc.
  • Often less efficient (8 hours for 20 minutes of speech)
  • Not appealing to everyone

– Single speaker rate: 7 – 15 min. of recorded speech per hour
  • Oral annotations with no written transcription
    • Enter comments in ‘notes’ tier when necessary
Kwak’wala

• Simultaneous written annotation
  Enter free translation into ELAN
  Mark sensitive material (index tier)
  Handwrite Livescribe fieldnotes:
    Kwak’wala transcription
    Additional comments, analysis, observations
• Rate: 7 – 15 min > 1 hour oral annotation
• Products:
  ELAN transcript file with English phrasal
  translation
  Additional recordings of transcription
  sessions
Oral annotation: Fieldnotes with Livescribe

- Records and creates automatic PDFs with time-aligned embedded audio
- Ideal for oral annotation sessions and post-trip processing
Oral annotation: Fieldnotes with Livescribe
Kwak’wala: post-trip

• Transfer Kwak’wala transcriptions from fieldnotes to primary ELAN transcription
• Track notes, comments, variation and other observations in separate tier
• Linguistic Research Apprentice Program (LRAP, UC Berkeley): undergraduate assistants
• Archive .wav, (.mov), .eaf and Pencast PDF files
• As needed: segment transcription sessions in ELAN
Discussion
Benefits

• Reduces transcription bottleneck
• Distribution of workload
  – Allows people less-trained in linguistics or less familiar with a particular language to successfully participate in transcription
• Opportunities for capacity development
  – Community language workers, L2 learners
  – Junior members of documentation team (i.e. undergraduates)
Challenges

• Proliferation of data (Cox 2013):
  – Multiple recordings of transcription sessions, relating to original recording:
    • How to treat these? As primary source or secondary source? Transcribe?
  – For now:
    • Original conversation is priority.
    • Secondary transcription session: chunk to enable navigation
    • Notes in ELAN transcript of original file regarding differences
Summary

• Extreme endangerment - time of elder speakers is a precious and limited commodity.

• Method increases accessibility of recorded conversation.

• Additional time spent on metadata and data management saves time elsewhere.

• Additional discoveries are possible through secondary recordings.
Kihchi-Marsii! Gilakas’la!

Daisy Rosenblum
drosemblum@umail.ucsb.edu

Olivia N. Sammons
sammons@ualberta.ca
References


