The Language

North-westernmost member of the Uto-Aztecan family
4-500 first language speakers
Speakers reside mainly in reservation communities scattered across large area
Most speakers over the age of 60 years

Archived matter housed mainly at the University of California-Berkeley (Berkeley Language Lab, Survey of California and Other Indian Languages) and the University of Nevada Reno (Special Collections Department)

Recent publication of dictionary (Liljeblad, et al. 2012)
~800 lines of oral narrative (Thornes 2003)

Fowler and Liljeblad (1986)

The Source Materials

UNR special collections is home to the Liljeblad Collection
Ethnographic and linguistic materials spanning 5 decades

Pete Snapp, master storyteller, was in his 90s when he began working with linguist and folklorist Sven Liljeblad to document narratives in the McDermitt dialect

Approximately 46 narratives recorded (~12 hours of audio)
Only ¼ of these carry any annotation (transcription and free translation)

Only the handwritten notes contain morphological analysis Astrid Liljeblad typed T & T for 12 narratives

Sven & Astrid Liljeblad
Pocatello, Idaho
1988
(photo courtesy of Michael J. P. Nichols)

The Problem

Related materials reside in disparate areas of the archive
No catalogued link between, say, Box 12 and Box 114 of the Liljeblad Collection

Cultural and Linguistic Interest of the materials

Pete Snapp was a master storyteller; high level of historical and cultural interest in content
Side conversations present further opportunity for analysis
Only known (although accidental) recording of child language
Informs discourse-grammatical features—use of clause-chaining, range of pronominal forms
Mapping of intonation units onto Liljeblad’s syntactic units is not always exact
Treatment of clitics appears true to notion of grammatical, not phonological, word criteria

The Project

Catalogue and pull together disparate archival materials
Use EUDICO Linguistic Annotator (ELAN) to align Pete’s speech with Sven’s annotations
Re-archive the digital output of the process

The Plan

→ Render the output searchable
→ Add tiers of annotation/analysis
→ Produce glossary, add to digital lexicon
→ Provide accessible version of materials to community
→ Complete annotations for entire corpus
→ Stay true to original documentation

The Process

Scan typewritten materials
Transfer and carefully edit as Word or Rich Text Document
Copy and paste line by line from Word document (to avoid error cascade)
Separate individual narratives into distinct audio files
Time align existing annotations (T & T) to audio

Additional tiers (thus far):
→ Intonation Units
→ Other speakers
→ Background
→ Collector footnotes

Contact

catherinef@uca.edu
timthornes@boisestate.edu

References