Making Field Recordings Useful: Online Tools for Accessing a Corpus of Recordings

Edwin Ko (Georgetown University)
Catherine O’Connor (Boston University)
Ethan Rimdzius (Boston University)
Elodie Paquette (Boston University)

Symposium on American Indian Languages
Friday, April 17th, 2015
Overview

❖ Background

❖ Northern Pomo

❖ Northern Pomo Language Tools

❖ Design and Technical Decisions

❖ Open-Source Template of these Tools
Motivation

- Many field recordings and field notes collected over 25 years on a language now dormant...
- Desire to make these resources accessible to the community of the descendants
- Language revitalization is very difficult...
- Ultimately, we decided to place on a website...
Goals

To make these resources...

❖ Widely accessible
  - Access by phones, tablets, or computers.

❖ Easily accessible
  - Structure the resources
  - Retrieve relevant materials

❖ Useful for many purposes
The process...

1. Creating Mobile apps
2. Creating Online Tools

Making them into open-source templates
The Pomoan Languages in Northern California (Kroeber, 1925)
Edna Campbell Guerrero  Elenor Stevenson Gonzales
Situation today...

- Currently no living fluent speakers.
- But, there is still a community that identifies with the language.
- EGIDS Level 9: Dormant:
  “The language serves as a reminder of heritage identity for an ethnic community, but no one has more than symbolic proficiency.”
Northern Pomo Language Tools
Exploring a language may be overwhelming...

Variety of different entryways into the language.

Start top-down, or create your own path.
northern pomo language tools

how to use the website
about the language
links to mobile apps
sounds and letters
everyday expressions
talking dictionary
phrasicon
stories and texts
contact
Some users may want to be oriented about the different sections of the website.
About the Language

Northern Pomo, one of seven distinct Pomo languages, was spoken in Northern California for thousands of years. Because of conquest and colonization by European settlers and government language policies, the language is no longer fluently spoken. This website is part of a project to archive and make available materials on the Northern Pomo language to support revitalization efforts.

These materials include voice recordings of the late Edna Campbell Guerrero and Elenor Stevenson Gonzales (pictured below), both native speakers of Northern Pomo. These materials would not exist without their patience, dedication, and hard work. The recordings were made by Catherine O’Connor between 1979 and 2005.

Edna Campbell Guerrero lived in Potter Valley until her death in 1995. Her father’s mother was from Sherwood Valley, in a band called shiwiwo toom pookim. Her mother had relatives from all over the area, including Ukiah and to the south. Elenor Stevenson Gonzales was an elder of the Pinoleville Rancheria. In her later years she lived in Santa Rosa with her daughter and grandchildren until her death in 2005.
Links to Mobile App

As a first step towards learning the Northern Pomo language, you might like to look at a mobile app that will introduce you to 85 Northern Pomo words with pictures. The app contains sound files by fluent speakers Edna Campbell Guerrero and Eleanor Stevenson Gonzales. The words can be studied within one of 7 categories (food, nature, people, places, animals, in the home, traditional objects) or across the entire set of 85 words.

Download

Clicking on the images below will bring you to the download page of the app.

- iOS App
  Currently, only the iPad version of the app is available. The iPhone version soon to come!

- Android App
  The Android version of the app is available and contains a quiz function as well.

➤ Links bringing you to the mobile apps.
Introductory word-learning app featuring 85 words.

Presented at an earlier talk today...
(Besides the mobile apps), these are the main components for incorporating field recordings...

northern pomo
language tools

how to use the website
about the language
links to mobile apps
sounds and letters
everyday expressions
talking dictionary
phrasicon
stories and texts
contact
northern pomo language tools

how to use the website
about the language
links to mobile apps

- sounds and letters
- everyday expressions
- talking dictionary
- phrasicon
- stories and texts
- contact
Learners may want to hear more about the sounds and how they’re written.

Matrix of sounds...

Clicking on the letter lets you hear the corresponding sound.

<table>
<thead>
<tr>
<th>?</th>
<th>a</th>
<th>a:</th>
<th>b</th>
<th>ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>ch'</td>
<td>d</td>
<td>e</td>
<td>e:</td>
<td>h</td>
</tr>
<tr>
<td>i</td>
<td>i:</td>
<td>j</td>
<td>k</td>
<td>k'</td>
</tr>
<tr>
<td>kʰ</td>
<td>l</td>
<td>m</td>
<td>n</td>
<td>o</td>
</tr>
<tr>
<td>o:</td>
<td>p</td>
<td>p'</td>
<td>pʰ</td>
<td>s</td>
</tr>
<tr>
<td>sh</td>
<td>t</td>
<td>t'</td>
<td>tʰ</td>
<td>t̚</td>
</tr>
<tr>
<td>t̚'</td>
<td>t̚ʰ</td>
<td>ts</td>
<td>ts'</td>
<td>u</td>
</tr>
<tr>
<td>u:</td>
<td>w</td>
<td>x</td>
<td>y</td>
<td></td>
</tr>
</tbody>
</table>
- Videos training learners linguistically significant sounds.
- Recordings of native speakers saying the similar words with contrasting sounds.
<table>
<thead>
<tr>
<th>ha</th>
<th>mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>xa</td>
<td>water</td>
</tr>
</tbody>
</table>
northern pomo language tools

how to use the website
about the language
links to mobile apps
sounds and letters
everyday expressions
talking dictionary
phrasicon
stories and texts
contact
Learners may want to learn expressions used on a day to day basis.

- "What is this?" "That is..."
- "Whatcha doing?"

On this page you will find links to short videos explaining some common expressions.
- Short videos that break down the expressions.
- Learners get to hear a fluent speaker saying that expression.
- Encourage repetition of the utterance.
How are you?
northern pomo language tools

how to use the website
about the language
links to mobile apps
sounds and letters
everyday expressions
talking dictionary
phrasicon

stories and texts
contact
Small set of around ten monologues by Edna Campbell Guerrero, recorded 1963.

Include narratives and monologues about cultural practices and personal experiences.
One approach...

- Videos of narratives with Northern Pomo and English translation.

- “Lyrics” appear on screen as the speaker is telling the story.

- There is also an accompanying text below that can be printed out.
nordern pomol
guages tools
how to use the website
about the language
links to mobile apps
sounds and letters
everyday expressions
talking dictionary
phrasicon
stories and texts
contact
Talking Dictionary...
Northern Pomo Online Talking Dictionary

English to Northern Pomo

a b c d e f g h i j k l m n o p q r s t u v w x y z

Northern Pomo to English

? b ch ch’ d h j k k‘ l m n p p’ p s sh t t’ t‘ t‘’ t‘’’ ts ts’ t‘ x y

You can type in Northern Pomo letters by clicking on the arrow to the right of the submit button.
Dictionary entries displayed in tabular format.

Link from Talking Dictionary to Phrasicon entries containing that head word.

Recordings and initials of the multiple speakers.
Entries in the Phrasicon in interlinear gloss format.

- First layer: sentence or phrase in Northern Pomo
- Second layer: morpheme
- Third layer: English gloss
- Fourth layer: English translation

Linkage between Talking Dictionary and Phrasicon works both ways.
Design and Technical Decisions
“The perfect dictionary is one in which you can find the thing you are looking for preferably in the very first place you look.” – Mary Haas

- A learner may not know what’s in the dictionary...
- What are some ways to help the user get what they want?
Scenario 1:

- Learner wants to know what the Northern Pomo word is for “brother”...
- In Northern Pomo, the language distinguishes between “older brother”, “younger brother”...there is no one word for “brother”
Organizing search term...

- Searching for "brother"...
Scenario 2:

- Learner wants to know the Northern Pomo word for “dogs”...
- In general, Northern Pomo does not distinguish between singular and plural nouns.
We used a technique, called a *stemmer*, that can look for the singular rather than the plural. Searching for “dogs” will return no result...
An option to try an ‘Extended Search’ that stems the search query. Searches again with “dog”...
Persistent Search Conundrum:

• User doesn’t know what to search in Northern Pomo, they will be searching in English.
• User search for the word “old”...
  • Should “older” be returned?
• But if the user searches for the word “man”
  • Should “manner” be returned?
Search rules...

• Search for “old” should return “older”
• Search for “man” should not return “manner”
• How to get around this issue?
• We implemented an ad-hoc search rule...
There are many persistent problems...

- Perhaps we should be talking about these kinds of issues as a field...
- Resource design principles and best practices.
Open-Source Templates
Language XYZ

- how to use the website
- about the language
- links to mobile apps
- sounds and letters
- everyday expressions
- talking dictionary
- phrasicon
- stories and texts
- contact
Current beta tests...

- Medumba, a Grassfields Bantoid language, spoken in Cameroon.
- Lachixío Zapotec, an indigenous language of Oaxaca, Mexico, with Mark Sicoli.
Our suggestion...

- Your team have a “coding competent” member who’s not afraid of looking into the code.
- Do not necessarily need to know to program, but should be able to follow directions in the documentation.
- Knowing how to program is helpful.
Where to download...


Github: eddersko
Acknowledgments:

Our deepest thanks to Edna Campbell Guerrero and Elenor Stevenson Gonzales for their many years of patient generosity and good humor. Thanks to the Tribal Historic Preservation Officers of several Northern Pomo Rancherias, and to Erika Estrada Carson.

The current project is supported in part by a grant from the NEH/NSF Documenting Endangered Languages program (FN-50107-12) to Catherine O'Connor, which is gratefully acknowledged. Previous grants from the National Science Foundation and National Endowment for the Humanities supported data collection in the 1990s. Views expressed in this material are the work of the authors and do not necessarily reflect the views of the U.S. National Science Foundation or the National Endowment for the Humanities.