Language Documentation in Nepal: Between the Micro- and the Macro

- Kristine Hildebrandt
- Shunfu Hu
- Tiffany Downing
- Andrea Fyffe
- Cassidy Jacobson
- Kanchan Karki
- Alex Taitt
• Linguistics: the scientific study of human language, concerned with describing & explaining the nature (and origins) of human language, how languages can vary (through space and time), and factors that can promote or inhibit the survival of a language across generations

• When linguists study ‘grammar’ they are concerned with what speakers actually do (vs. what they are taught they ‘should’ do through prescriptive norms), and ‘grammar’ can include more than just words and sentences: speech sounds, meaning, language attitudes, social variation, non-spoken communication …
The over-arching goal of this project is **language documentation**: the creation of a comprehensive record of the language structure & practices of a speech community.

We are interested in creating a full & complete a record of the languages, both for posterity (should they die), and for possible language revitalization (if there is a community initiative for revival).

This record also makes that language system available for linguistic analysis, to contribute to a fuller understanding of human language in other dimensions, including spatial (geography) social (sociology) and cultural (narrative and text studies).

The presentations and demonstrations today illustrate these kinds of activities.
The Project: Documenting the Languages of Manang for Local and International Impact

- For the past three years a group of faculty and students from SIUE and other have collaborated on documenting four languages of Manang, Nepal
- This project is jointly funded by the National Science Foundation (DEL 1149639) and the SIUE Undergraduate Research and Creative Activities Program

Our Nepal field site teams in 2012 & 2013

www.mananglanguages.org
The Languages: Four Tibeto-Burman Languages spoken in the Manang District, Nepal

- **Nar & Phu** spoken to north-east (600 spkrs total)
- **Nyeshangte/Manange** spoken in northern & central VDC’s (3,000-5,000 spkrs)
- **Manang-Gurung** in southern & central VDC’s (2,000 spkrs)
- **Gyalsumdo** around Tal & Chame (< 300 spkrs)
Today’s Panel

• We hope to show how our work on this project moves between the large-scale and the small-scale in both the questions that seek to answer, as well as the data that we include in our investigations

1. Shunfu Hu and Kanchan Karki (Geography): A spatial perspective of these four languages: The Atlas

2. Tiffany Downing and Cassidy Jacobsen: Reaching out to different communities through sharing of text and narrative samples

3. Andrea Fyffe: Language attitudes, language practices: A brief comparison of similarities and differences in different communities

4. Alex Taitt and Kristine Hildebrandt: Close-zoom: Investigating the tone systems of the Manang Languages

• Closure & Q/A
Geography of Manang District, Nepal
Naar village
Language Atlas
Cloud-based data storage and access

- Google Maps API (digital maps and satellite imagery)
- Yahoo! Flickr API (digital photographs)
- YouTube API (digital video)

Web Host Service (Server)

Network connections to devices:
- Desksops
- Laptops
- Tablets
Development environment:

- HTML5
- XHTML
- JavaScript
- XML
- jQuery
Initial launch of the atlas
The user can select a particular village, for instance, Danakyu, from the drop-down list and then click the Search button to show all the sociolinguistic interviews conducted in that village, three in this example. The user can click on an icon on the map to get the picture of the house where the interview took place and watch the video of the interview. The user can also download the acoustic data for the interview by clicking on the link.
The user can select one particular language (e.g., Manang Gurung) to see where this language is being used in all the villages in Manang District.
Movements of Texts

• Importance of texts
  o Specifically Gyalsumdo and Gurung
  o Allows us to share information about culture and history
  o Texts show aspects of their daily life
  o Share opinions about their language

• How stories are shared
  o Youtube
  o SHANTI
• Youtube gives the general public a way to view these videos
• By using ELAN software and Windows Movie Editor we were able to subtitle the videos so English speakers are able hear the language and understand each person’s story.
• Metadata gives some basic information about the speaker, time and place, and the technology used.
https://www.youtube.com/watch?v=uBq_IpbMYI8
SHANTI

- Tibetan and Himalayan Library
  - SHANTI audio video collection
- Different audience
- Subtitling
- Metadata

http://mediabase.shanti.virginia.edu/collection/gyalsumdo-project
SHANTI

A Gyalsumdo Man Explains His Religious Beliefs

<table>
<thead>
<tr>
<th>SAB</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAB</td>
<td>na minda ogji lo fo tere People have called me Akikung Lama.</td>
</tr>
<tr>
<td>SAB</td>
<td>ni gyalsumdo mi dnoi If a man dies in our Gyalsumdo community,</td>
</tr>
<tr>
<td>SAB</td>
<td>dlo, fei'ms lapasami if someone dies,</td>
</tr>
<tr>
<td>SAB</td>
<td>nba me dhene bno phep, ji kere we should call the Lama,</td>
</tr>
<tr>
<td>SAB</td>
<td>bno phep, ji kere We should call the Lama.</td>
</tr>
<tr>
<td>SAB</td>
<td>lamu ki phowa kip kip kere The Lama helps people cross to the heaven.</td>
</tr>
</tbody>
</table>

http://mediabase.shanti.virginia.edu/node/2123/transcript
Sociolinguistics

• Sociological Interviews
  – It can be used to explore how a respondent feels about a particular topic. This provides the opportunity to gather a greater depth of information.
  – It can allow the building of relationships with other people, from different cultures.
  – A sociological survey allows the researcher to ask personal questions to the interviewee, which once again allows for a more in depth conversation with the individual.
The Bigger Idea

• We are trying to gather personal opinions and information based on the language system of Nepal.

• The responses to these questions can give insight on the past and present practices and attitudes.

• And the responses can shed light on the potential future prospects for these languages given the economic and geographic changes happening in Manang.

• The following slides show some responses to selected questions for selected languages.
What Language Do You Use With Your Children?

- Manange: 32%
- Nepali: 21%
- Manange + Nepali: 16%
- I speak Manange, They respond Nepali: 16%
- Nepali + English: 5%
- Tibetan + English: 5%
Gurung

Is Your Mother Tongue Important To Your Culture?
Gurung

Agree

97%

Disagree

3%
Gyalsumdo

Should Children Be Able To Decide Their Language Preference?
Gyalsumdo

- No: 73%
- No Opinion: 20%
- Yes: 7%
Zooming In To The Grammar!

- A microscopic level of these languages
- ZOOM in to view each word under a “microscope”
- Our most linguistic-specific focus in the overall project (so far...)

![Image of two people working on a laptop]
Tone Systems

- Most people think of this when they hear about a ‘tone language’ (Mandarin Chinese)
- But because tone is still emerging in the Manang languages, it is very subtle and needs to be studied in terms of many different sound (acoustic) cues

<table>
<thead>
<tr>
<th>Chinese Character</th>
<th>Tone symbol</th>
<th>Tone description</th>
<th>English gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>媽</td>
<td>↓</td>
<td>high level</td>
<td>‘mother’</td>
</tr>
<tr>
<td>麻</td>
<td>↑</td>
<td>high rising</td>
<td>‘hemp’</td>
</tr>
<tr>
<td>馬</td>
<td>↘</td>
<td>falling rising</td>
<td>‘horse’</td>
</tr>
<tr>
<td>罵</td>
<td>↓</td>
<td>high falling</td>
<td>‘scold’</td>
</tr>
</tbody>
</table>
## Gyalsumdo Tone

<table>
<thead>
<tr>
<th>Word + meaning</th>
<th>Tone</th>
</tr>
</thead>
<tbody>
<tr>
<td>tʃik ‘one’</td>
<td>high</td>
</tr>
<tr>
<td>tʃʰi(ʃon) ‘to die’</td>
<td>high, with aspiration</td>
</tr>
<tr>
<td>ŋiŋ ‘field’</td>
<td>low, murmured</td>
</tr>
<tr>
<td>ʃi ‘four’</td>
<td>low &amp; falling, murmured</td>
</tr>
</tbody>
</table>
How do we examine this?

- Repetitive measurement of tiny details: intensity, pitch, jitter, duration, vowel formant ratios, consonant aspiration
- multiple measurements * 6 word repetitions * 65 words * 6 speakers = too much manual work!! 😞
Alternative—some automation

• Write computer scripts to automate these measurement processes
• Same word, same sound analysis software...
• **BUT:** added boundaries “tell the” script where to find needed data
• **RESULT:** measurements are faster and more reliable
Alternative—some automation

- Write computer scripts to automate these measurement processes
- Same word, same sound analysis software...
- **BUT:** added boundaries “tell the” script where to find needed data
- **RESULT:** measurements are faster and more reliable

```
# A TextGrid and sound file must be selected in the object list to proceed.
sound$ = selected$("Sound",1)
grid$ = selected$("TextGrid",1)

select TextGrid 'grid$
# ask the user for the tier number

form Calculate durations of labeled segments
comment Which tier of the TextGrid object would you like to analyse?
integer Tier 3
comment What is the name of your output file?
text Output ThMDurEGG1output.txt
endform

# check how many intervals there are in the selected tier:
numberOfIntervals = Get number of intervals... tier

#a = 0
# loop through all the intervals
for interval from 1 to numberOfIntervals
label$ = Get label of interval... tier interval
# if the interval has some text as a label, then calculate the duration.
if label$ <> ""
```
Data Analysis

• Next step: data dump into a spreadsheet for organization
• This makes statistical analysis easier & faster
What We’re Learning: Gurung Tone

Pitch Properties By Tone (Female #1)

Pitch Properties By Tone (Female #2)

Pitch Properties By Tone (Male #1)