

# 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

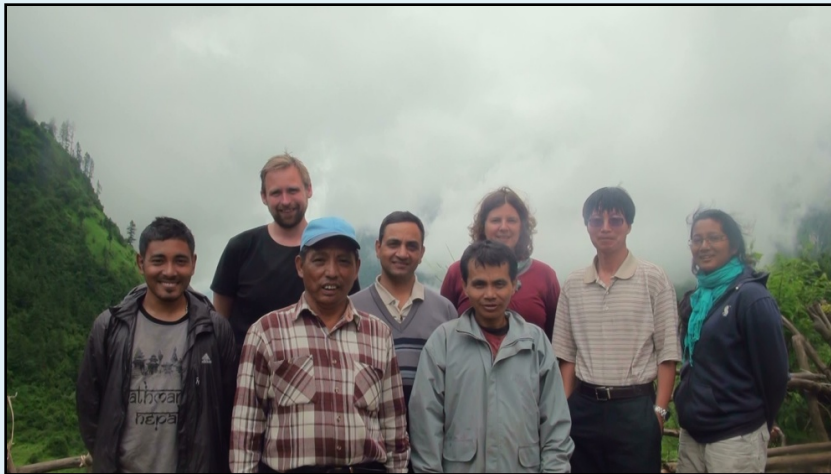


- We are interested in creating a full & complete a record of the languages, both for posterity (should they die), and for possible language revitalization (if there is a community initiative for revival)
- This record also makes that language system available for linguistic analysis, to contribute to 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



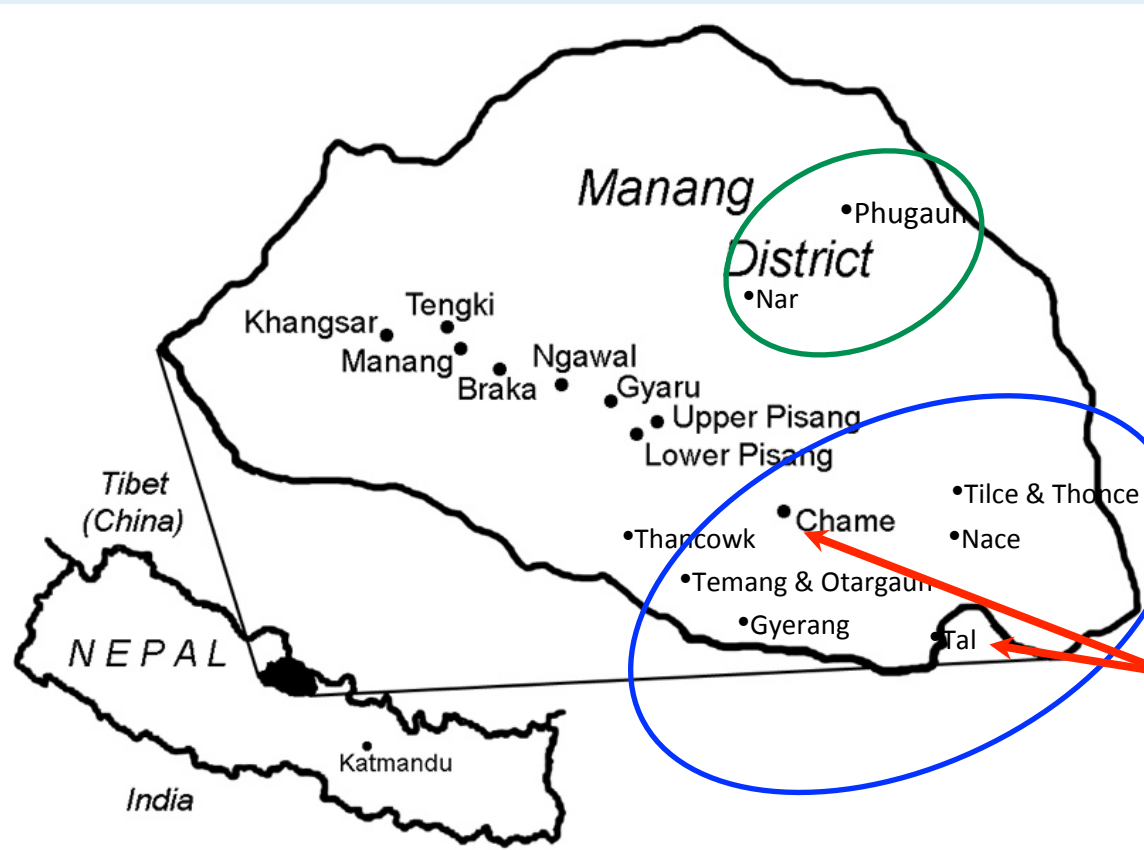
# The Project: Documenting the Languages of Manang for Local and International Impact

- 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*

# 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



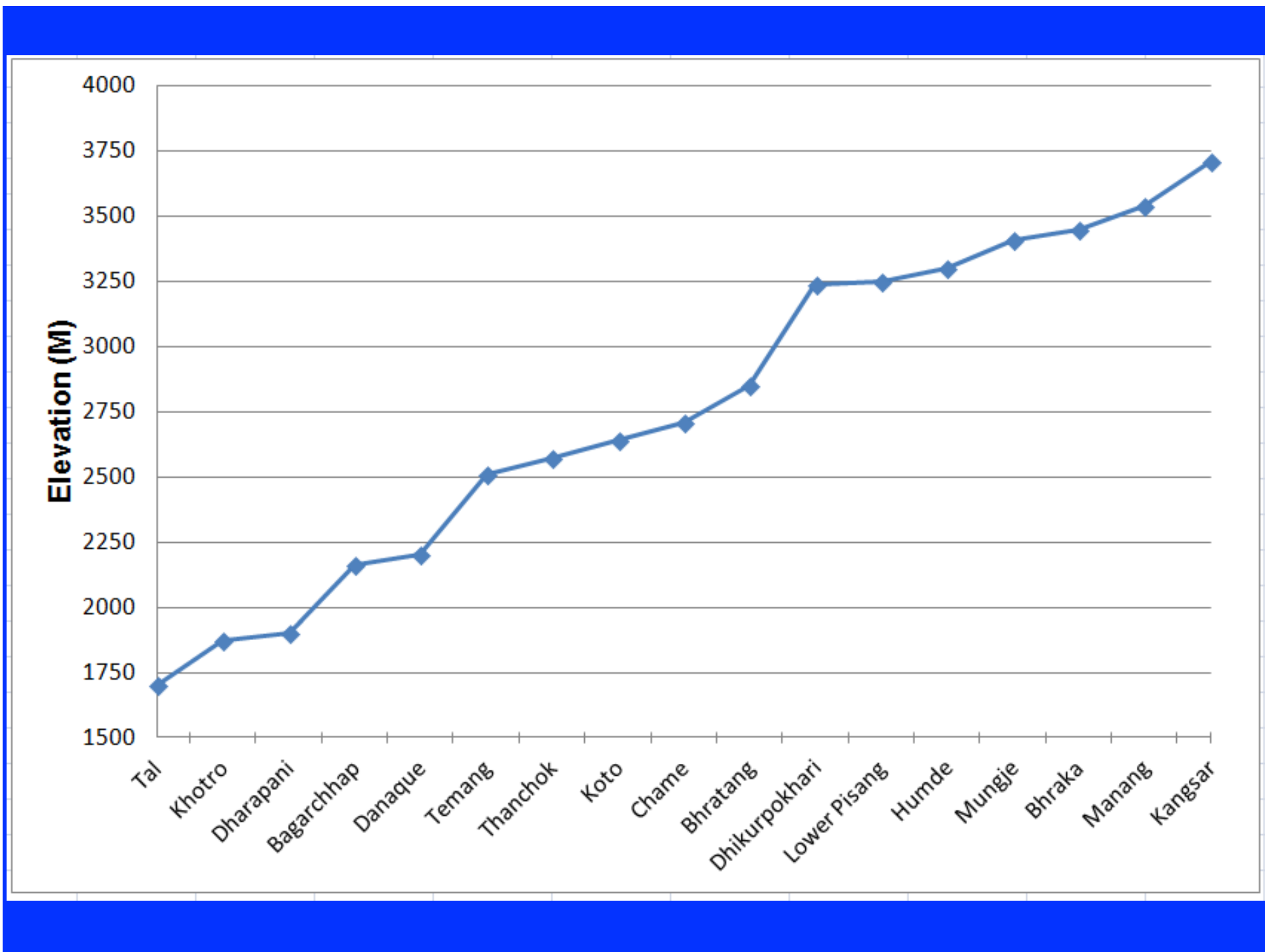






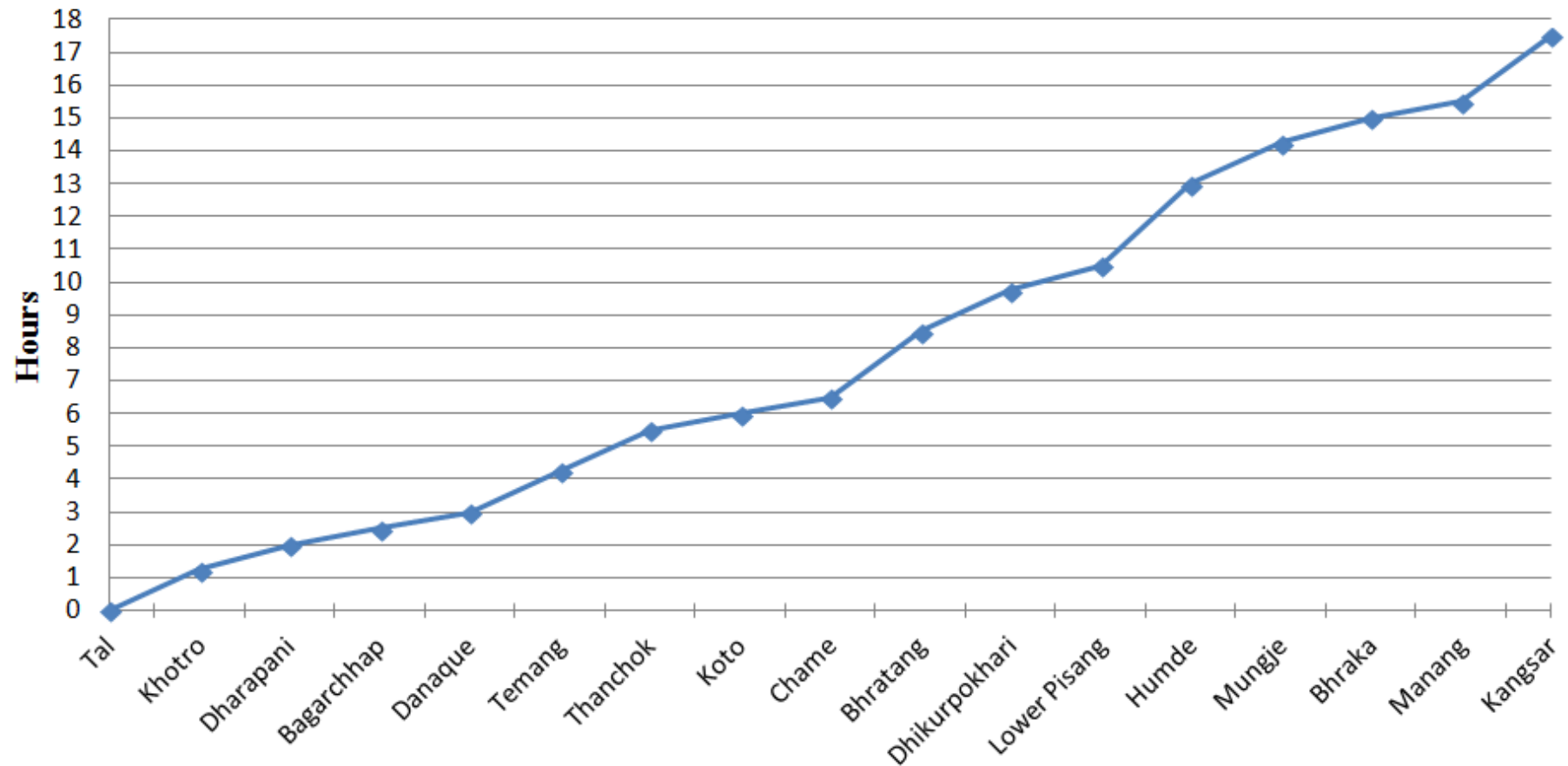








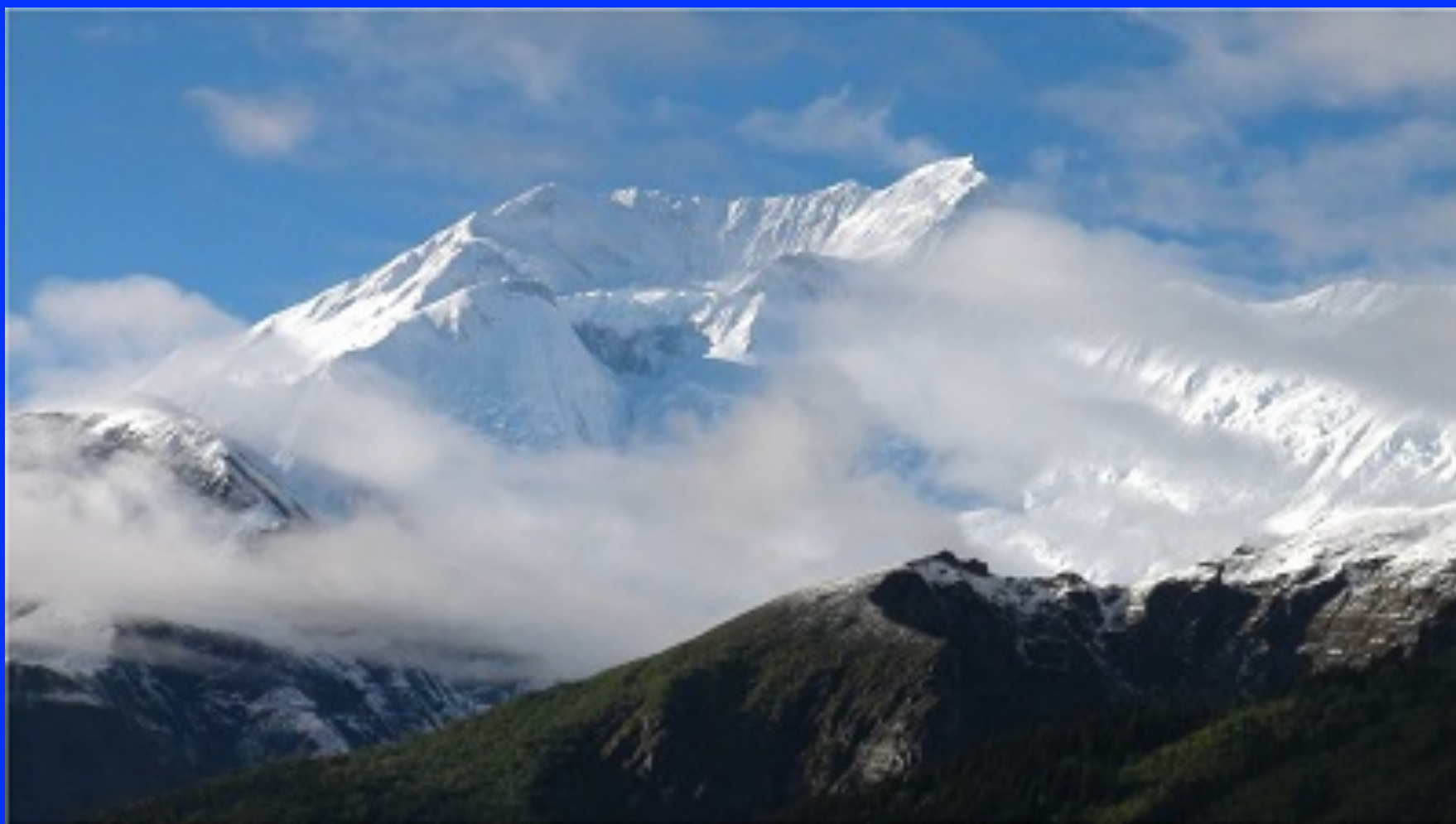
## Trekking Time from Tal Village





























Mani Walls





Naar village











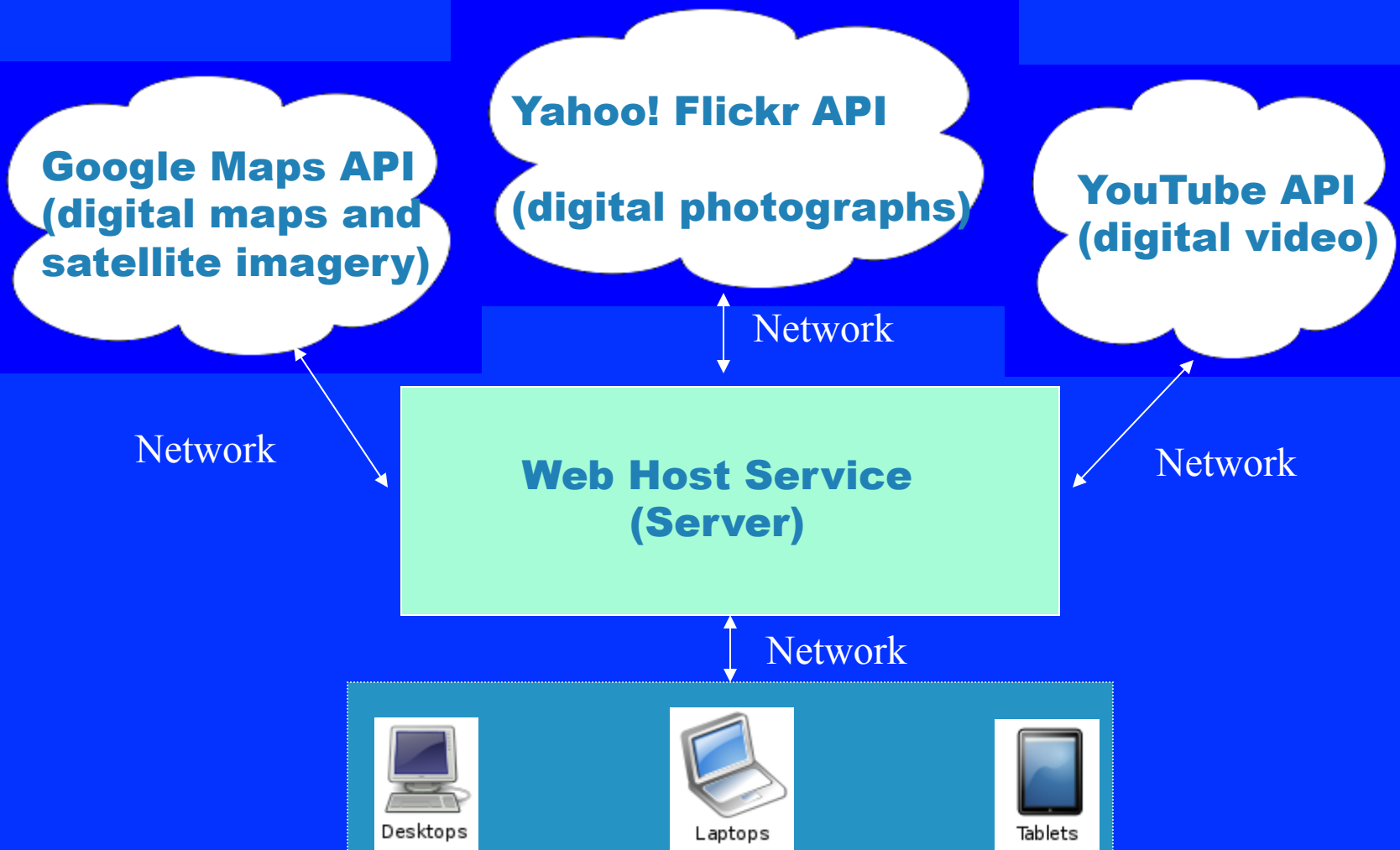






# Language Atlas

# Cloud-based data storage and access

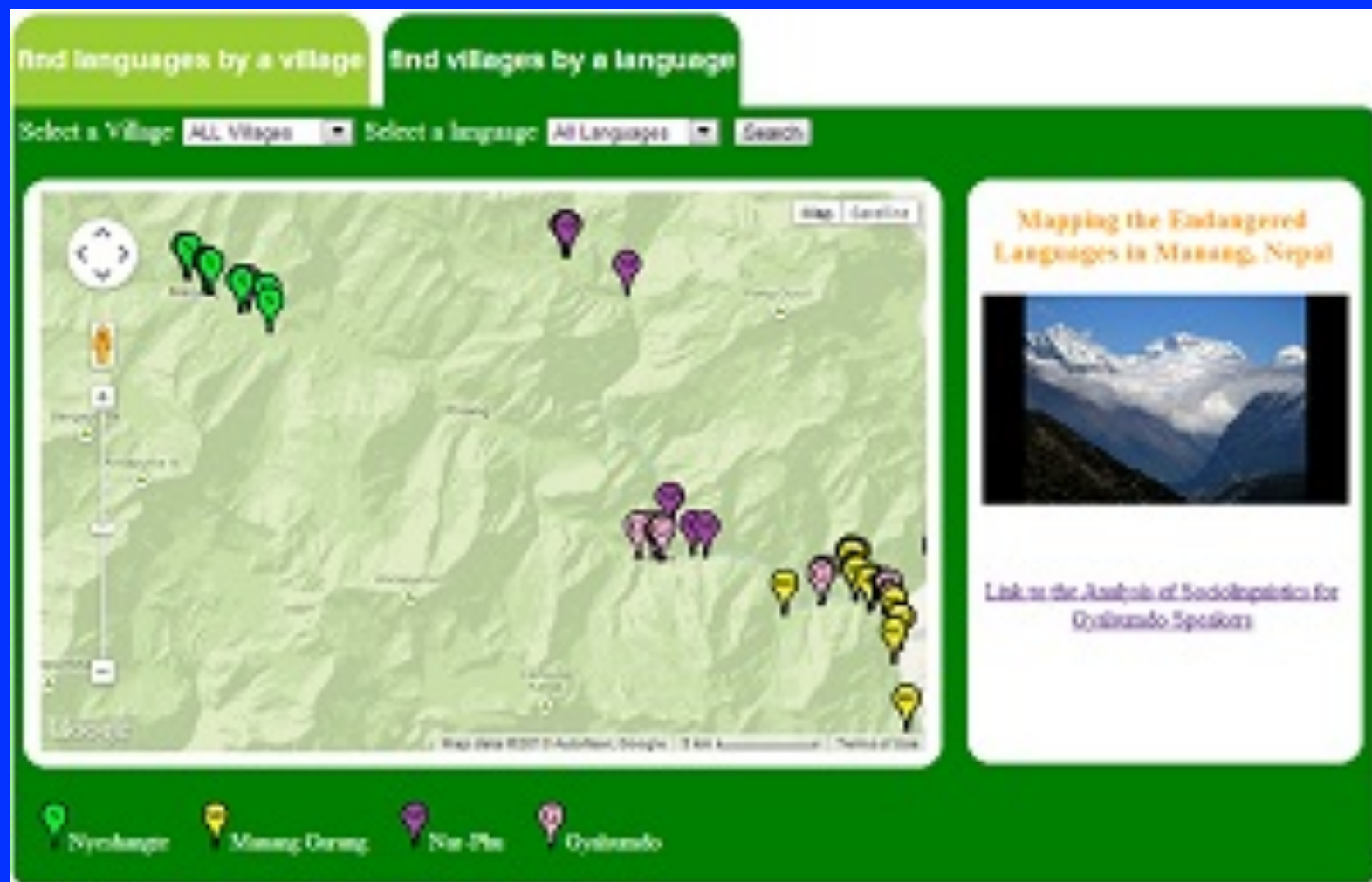




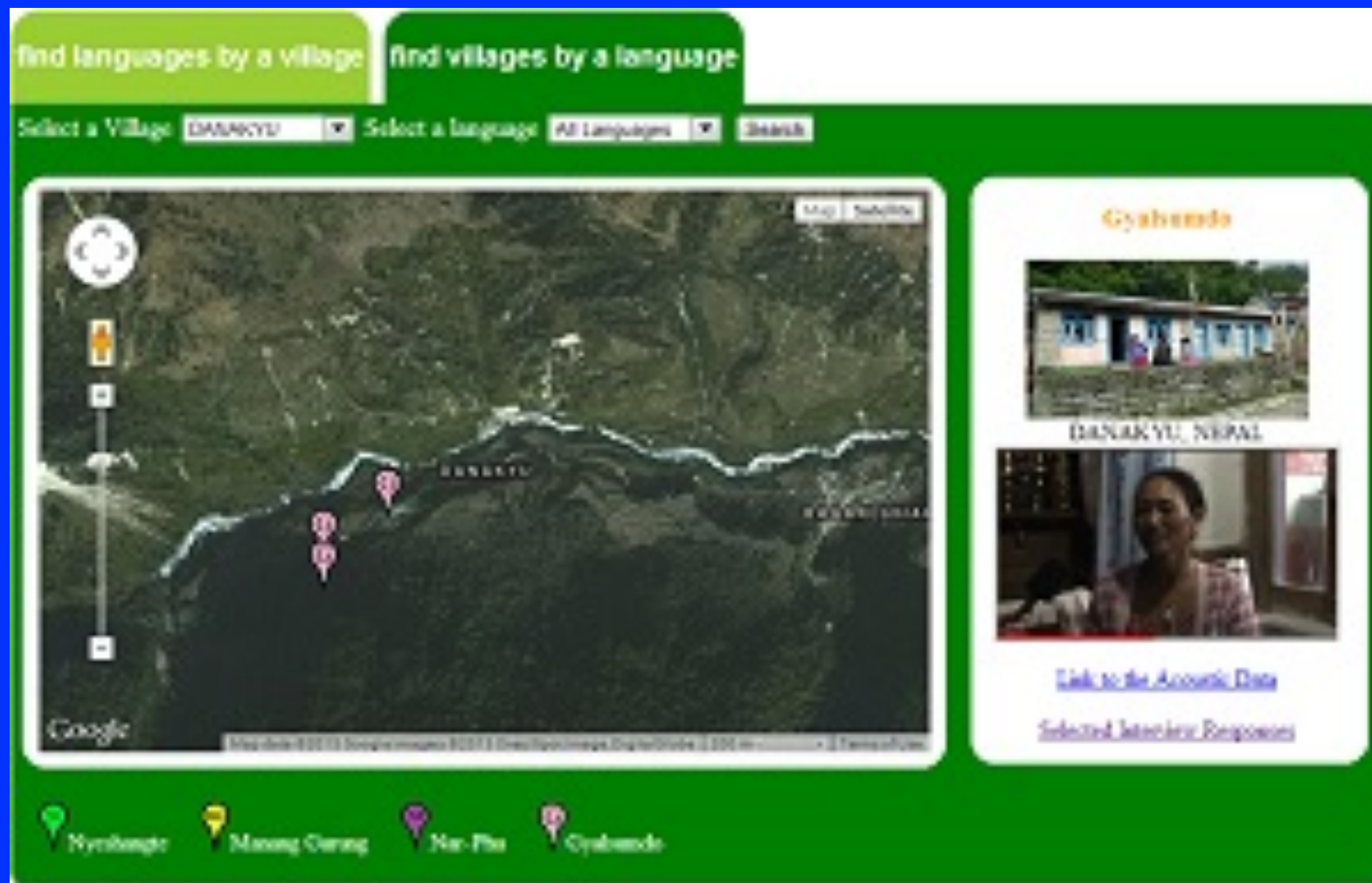


# Development environment:

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**
  - **Specifically Gyalsumdo and Gurung**
  - **Allows us to share information about culture and history**
  - **Texts show aspects of their daily life**
  - **Share opinions about their language**
- **How stories are shared**
  - **Youtube**
  - **SHANTI**

# Youtube

- 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.



# Youtube

[https://www.youtube.com/watch?v=uBq\\_IpbMYI8](https://www.youtube.com/watch?v=uBq_IpbMYI8)



# SHANTI

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

## Gyalsumdo Project

### Collection

- Manager: [Kristine Hildebrandt](#)
- 10 media item(s)

Gyalsumdo is a variety of Tibetan that is spoken in a cluster of villages in the lower Manang District of Nepal (Gandaki Zone, 22° 59' N; 84° 22' E). There is currently no ISO 639 code assigned to Gyalsumdo. Based on existing available literature, Gyalsumdo has very little prior documentation available on it aside from passing mention of Tibetan-speaking peoples in some Manang villages, and there is still question as to its specific placement/affiliation within lower levels of the Tibeto-Burman taxonomy. Based on comparative lexical information with Kyirong Tibetan, Nubri and Tsum (spoken in Gorkha District), Hildebrandt and Perry (2011) propose a distinct sub-grouping within Central Tibetan to which Gyalsumdo belongs, along with Kyirong and Nubri.

The goal of this collection is to assemble a variety of texts, representing different genres, from different speakers of Gyalsumdo who either had been born and raised in one of the three main villages affiliated with Gyalsumdo people (Chaame, Danakyu/Bagarchhap, Thonce), or else who had migrated to other regions of Nepal. As there currently is no agreed-upon orthographic representation for Gyalsumdo community internally, the representation of the language is in IPA and the translations include English and Nepali (the contact language).

These materials comprise a portion of a larger language documentation project taking place in Manang, Nepal (National Science Foundation DEL1149639 and Hans Rausing ELDP Small Grant SG0025). All materials were collected and are archived with speaker permission, and all recording practices confirm to standard Informed Consent protocols authorized by the institution hosting this language documentation project.

### Reference:

Hildebrandt, K.A. and J.J. Perry. 2011. Preliminary Notes on Gyalsumdo, an Undocumented Tibetan Variety in Manang District, Nepal. *Himalayan Linguistics* 10.1: 167-185.

### Media in Gyalsumdo Project



A Gyalsumdo Speaker  
Describes the History of  
His Family



A Gyalsumdo Woman  
Describes Her Wedding  
Ceremony



Ngima Pincho Lama,  
Danakyu village



A Gyalsumdo Speaker  
Describes His Village and  
Military Career



A Gyalsumdo Male  
Describes the Mineral  
Water Factory in Koto

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


# SHANTI

## A Gyalsumdo Man Explains His Religious Beliefs

Transcript Search

View Edit Transcript

Gyalsumdo English Nepali Speech bubbles One-line player Interactive transcript

 1:42

Previous Same line Next

SAB

ra minla atfju lab fjo tere  
People have called me Achjung Lama.  
मेरो नाम (लाई) आच्युङ् भन्ने गरेको छ।

SAB

ri galsumdo mi jinani  
If a man dies in our Gyalsumdo community,  
हाम्रो ग्यालसुम्दोमा मान्छे मरेमा

SAB

dha jefun lapanani  
,if someone dies,  
अब मर्यो भनेमा

SAB

filon me dhene lama phep ji kere  
we should call the Lama.  
मर्ने वित्तिकै लामा बोलाउनु पर्छ।

SAB

lama phep ji kere  
We should call the Lama.  
लामालाई बोलाउनु पर्छ।

SAB

lamaki phowa kjap lapa kere  
The Lama helps people cross to the heaven.  
लामाले आत्मालाई स्वर्ग तान्नु पर्छ।

<http://mediabase.shanti.virginia.edu/node/2123/transcript>



# Sociolinguistics

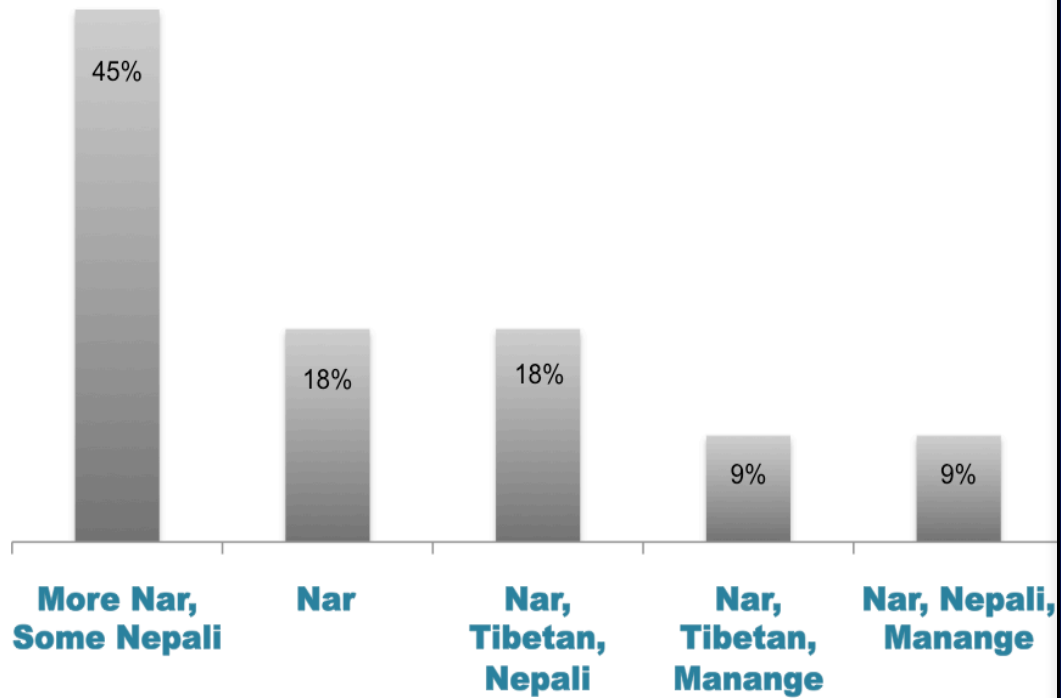
- Sociological Interviews
  - It can be used to explore how a respondent feels about a particular topic. This provides the opportunity to gather a greater depth of information.
  - It can allow the building of relationships with other people, from different cultures.
  - A sociological survey allows 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

# Nar-Phu

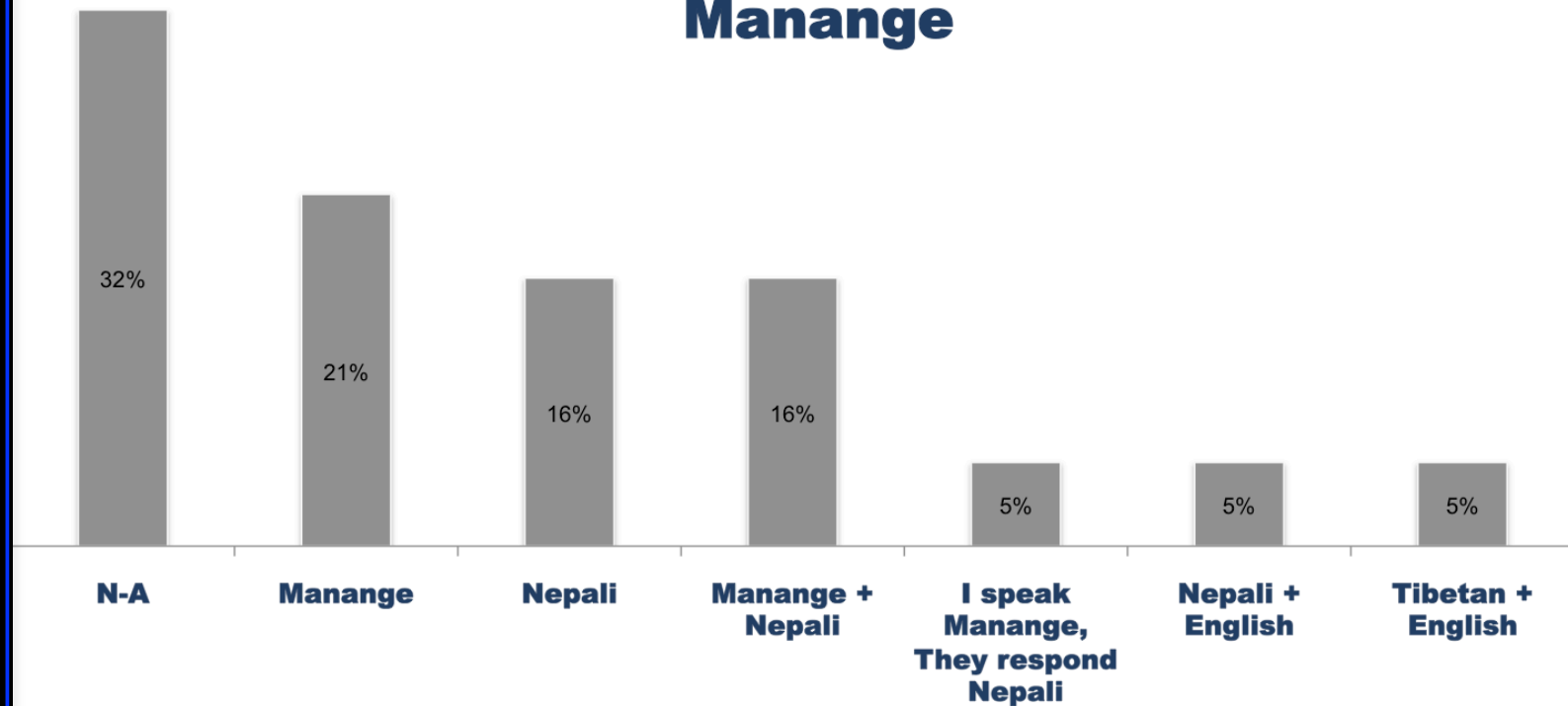
## Language Used In Daily Life Nar-Phu





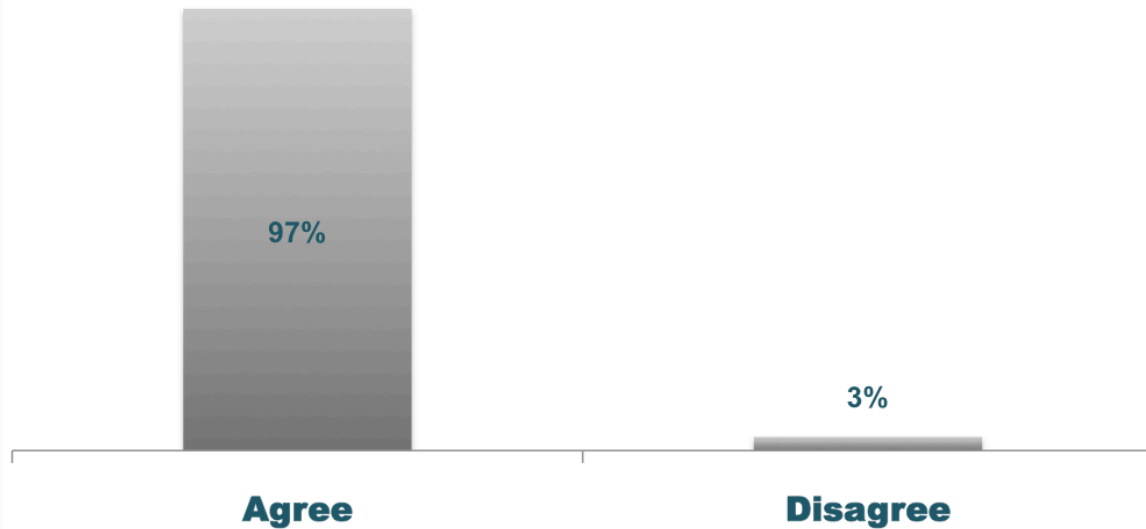
# Manange

## What Language Do You Use With Your Children? Manange



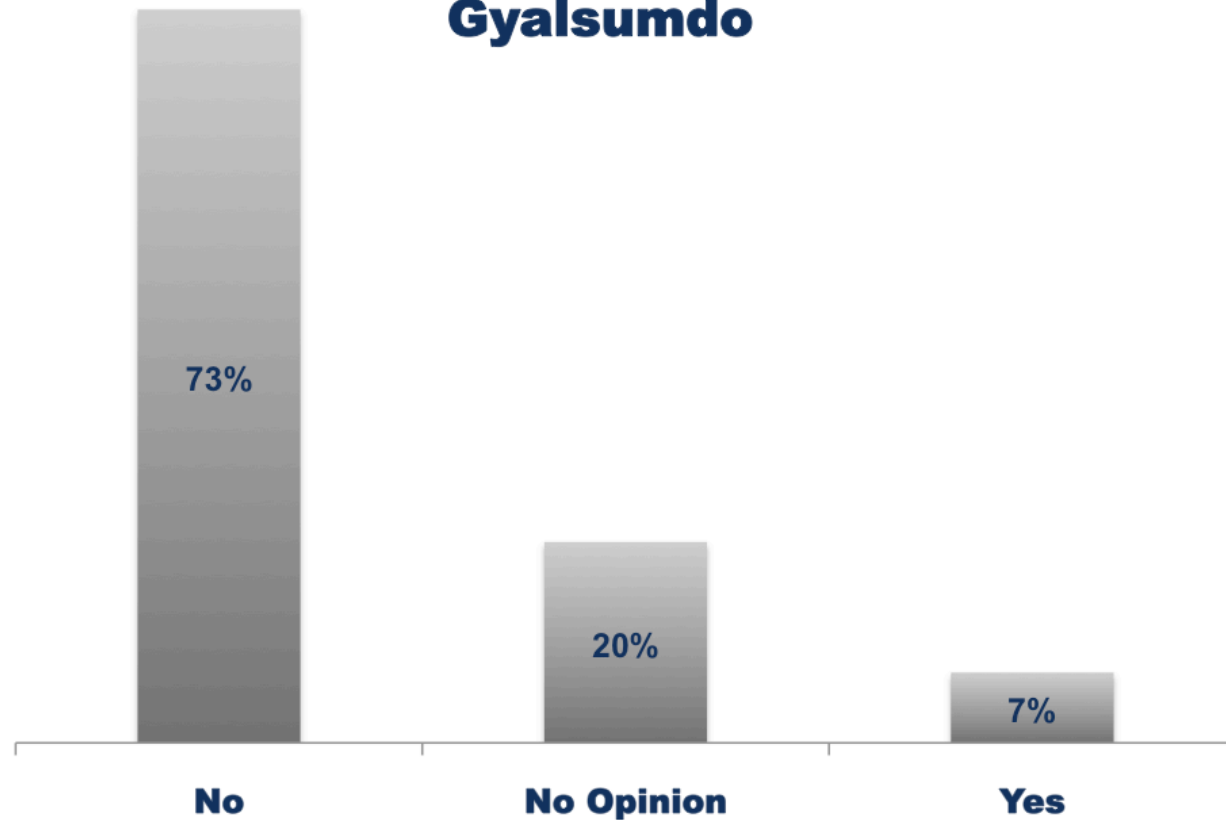
# Gurung

**Is Your Mother Tongue  
Important To Your Culture?  
Gurung**



# Gyalsumdo

**Should Children Be Able To  
Decide Their Language Preference?  
Gyalsumdo**





# 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...)



# Tone Systems

- Most people think of this when they hear about a 'tone language' (Mandarin Chinese)

STANDARD CHINESE <b>ma</b>			
Chinese Character	Tone symbol	Tone description	English gloss
媽	ˊ	high level	'mother'
麻	ˊˊ	high rising	'hemp'
馬	ˋˊ	falling rising	'horse'
罵	ˋˋ	high falling	'scold'

- 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
- Example from Gyalsumdo

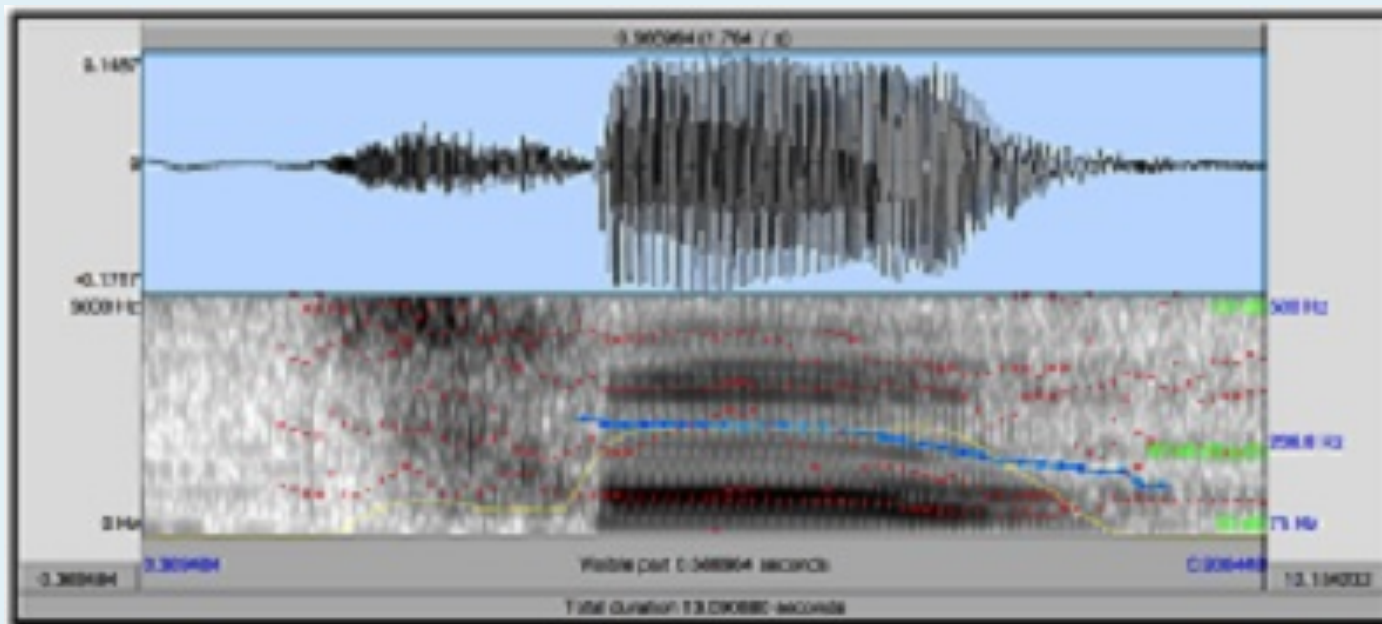
# Gyalsumdo Tone

Word + meaning	Tone
tʃik 'one'	high
tʃ <sup>h</sup> i(ʃon) 'to die'	high, with aspiration
ʃin 'field'	low, murmured
ʃi 'four'	low & falling, murmured



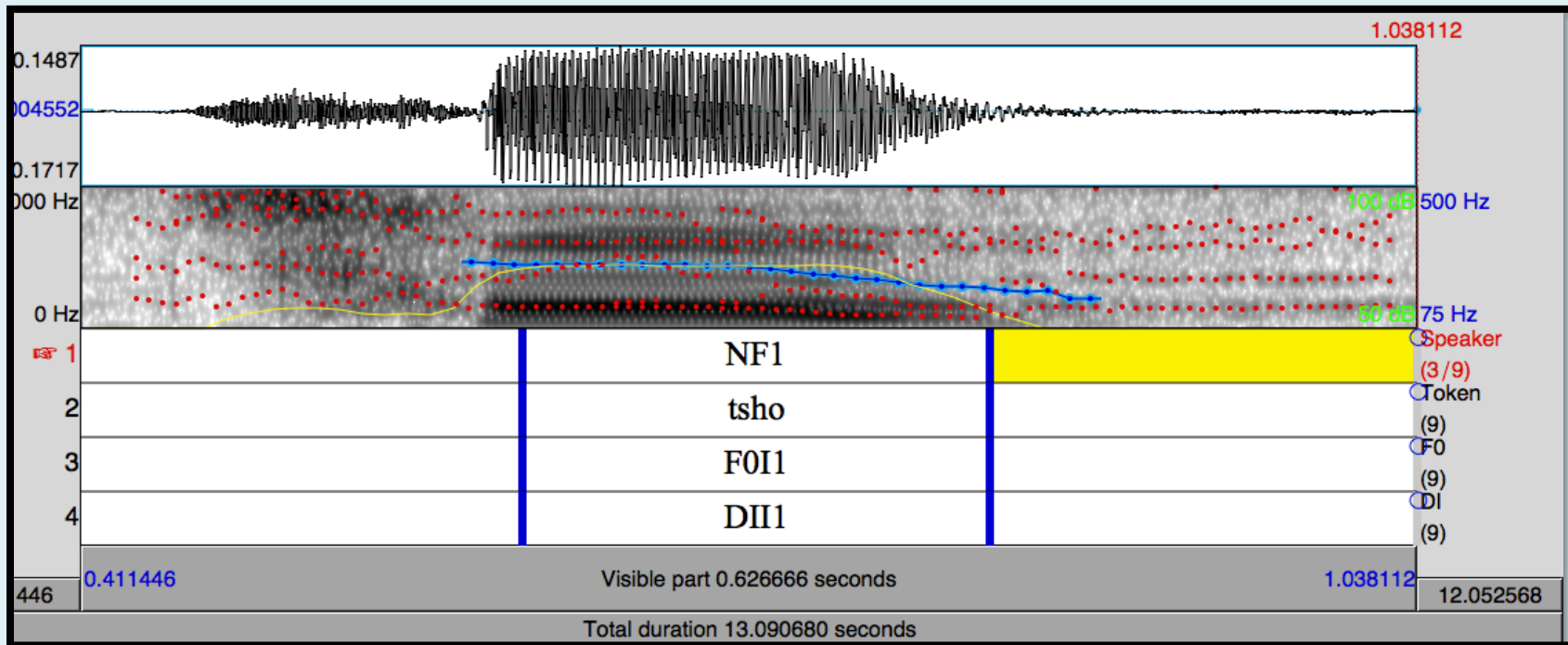
# How do we examine this?

- 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!!!** 😞



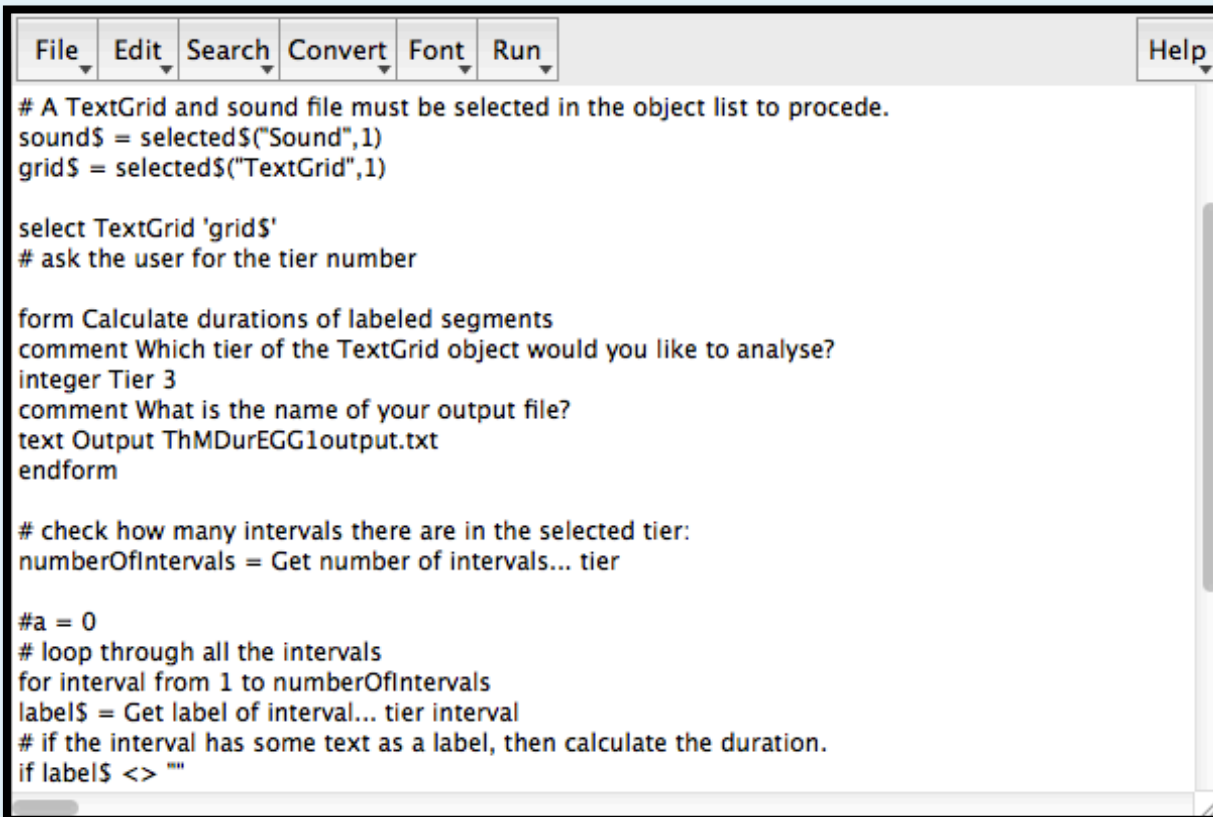
# Alternative—some automation

- 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



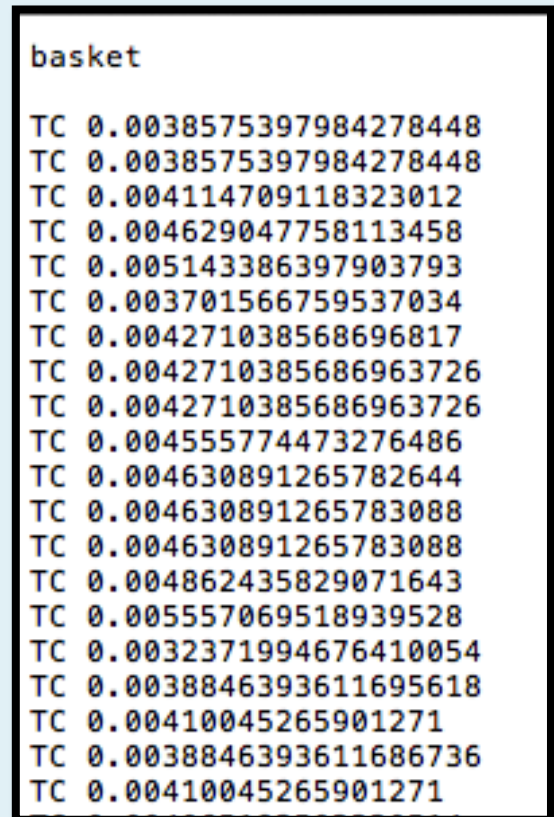
```
# A TextGrid and sound file must be selected in the object list to procede.
sound$ = selected$("Sound",1)
grid$ = selected$("TextGrid",1)

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$ <> ""
```



```
basket

TC 0.0038575397984278448
TC 0.0038575397984278448
TC 0.004114709118323012
TC 0.004629047758113458
TC 0.005143386397903793
TC 0.003701566759537034
TC 0.004271038568696817
TC 0.0042710385686963726
TC 0.0042710385686963726
TC 0.004555774473276486
TC 0.004630891265782644
TC 0.004630891265783088
TC 0.004630891265783088
TC 0.004630891265783088
TC 0.004862435829071643
TC 0.005557069518939528
TC 0.0032371994676410054
TC 0.0038846393611695618
TC 0.00410045265901271
TC 0.0038846393611686736
TC 0.00410045265901271
```



# Data Analysis

- Next step: data dump into a spreadsheet for organization
- This makes statistical analysis easier & faster

CE1141															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Tone_Tamu	Tone_KASK	Tone_MAN	INF	GLOSS	TC1W1	TC2W1	TC3W1	TC4W1	TC5W1	TC1W2	TC2W2	TC3W2	TC4W2	TC5W2
2	1	1	4	NF1	antler	0.0012	0.0011	0.0013	0.0017	0.0019	0.0013	0.0013	0.0014	0.0020	0.0022
3	1	2	3	NF1	basket	0.0011	0.0013	0.0013	0.0018	0.0022	0.0016	0.0017	0.0014	0.0018	0.0019
4	1	2	2	NF1	beam1	0.0012	0.0012	0.0019	0.0022	0.0023	0.0012	0.0013	0.0017	0.0022	0.0023
5	1	4	4	NF1	bee	0.0022	0.0023	0.0020	0.0022	0.0021	0.0018	0.0020	0.0021	0.0026	0.0027
6	2	1	1	NF1	bladder	0.0011	0.0013	0.0013	0.0018	0.0022	0.0016	0.0017	0.0014	0.0018	0.0019
7	2	2	2	NF1	blood	0.0014	0.0014	0.0017	0.0022	0.0024	0.0012	0.0012	0.0015	0.0021	0.0024
8	1	1	4	NF1	book	0.0021	0.0026	0.0026	0.0026	0.0029	0.0024	0.0026	0.0026	0.0028	0.0028
9	1	1	1	NF1	chest	0.0016	0.0013	0.0014	0.0016	0.0020	0.0016	0.0013	0.0013	0.0016	0.0018
10	2	2	1	NF1	chin	0.0019	0.0028	0.0043	0.0024	0.0026	0.0015	0.0022	0.0020	0.0029	0.0026
11	4	3	1	NF1	churn	0.0013	0.0015	0.0015	0.0015	0.0014	0.0013	0.0014	0.0014	0.0016	0.0016
12	1	4	4	NF1	cow	0.0017	0.0018	0.0019	0.0022	0.0026	0.0019	0.0021	0.0024	0.0028	0.0028
13	1	4	4	NF1	divide	0.0015	0.0015	0.0013	0.0015	0.0015	0.0013	0.0015	0.0015	0.0015	0.0016
14	1	2	1	NF1	do	0.0018	0.0019	0.0016	0.0016	0.0015	0.0018	0.0016	0.0013	0.0017	0.0016
15	1	3	1	NF1	drive	0.0013	0.0013	0.0016	0.0013	0.0014	0.0013	0.0012	0.0015	0.0016	0.0014
16	1	1	1	NF1	eat	0.0015	0.0019	0.0019	0.0019	0.0015	0.0019	0.0018	0.0021	0.0018	0.0021
17	3	2	2	NF1	eye	0.0019	0.0020	0.0019	0.0021	0.0021	0.0023	0.0020	0.0022	0.0022	0.0024
18	1	1	1	NF1	feces	0.0016	0.0017	0.0014	0.0020	0.0019	0.0017	0.0019	0.0013	0.0017	0.0018
19	1	2	1	NF1	feed	0.0010	0.0011	0.0013	0.0015	0.0013	0.0011	0.0015	0.0018	0.0016	0.0020
20	1	1	1	NF1	fire	0.0018	0.0017	0.0018	0.0021	0.0023	0.0017	0.0015	0.0018	0.0019	0.0026
21	1	1	1	NF1	flesh	0.0013	0.0018	0.0015	0.0023	0.0024	0.0015	0.0016	0.0020	0.0018	0.0023
22	1	3	3	NF1	forehead	0.0015	0.0014	0.0015	0.0015	0.0026	0.0021	0.0021	0.0021	0.0023	0.0028
23	4	2	2	NF1	full	0.0014	0.0013	0.0013	0.0014	0.0014	0.0013	0.0013	0.0016	0.0016	0.0016
24	1	2	1	NF1	give	0.0014	0.0016	0.0014	0.0017	0.0014	0.0017	0.0015	0.0015	0.0019	0.0024
25	1	4	1	NF1	go	0.0017	0.0018	0.0014	0.0018	0.0019	0.0022	0.0022	0.0023	0.0021	0.0021
26	1	1	1	NF1	grass	0.0010	0.0012	0.0013	0.0014	0.0016	0.0011	0.0012	0.0013	0.0015	0.0016
27	3	3	2	NF1	grind	0.0011	0.0012	0.0014	0.0013	0.0014	0.0011	0.0011	0.0013	0.0016	0.0017
28	1	1	1	NF1	ground	0.0021	0.0024	0.0014	0.0018	0.0026	0.0011	0.0013	0.0014	0.0018	0.0020

# What We're Learning: Gurung Tone

