Field Methods (LING404): Workflow

Most of our work this semester will follow a regular pattern of 'elicitation – transcription – dictionary editing'. This document lays out how, in detail, our work-week will proceed.

1.   Broad Overview of Workflow

1.1   Class Sessions

Each class session is a collective interview with our language consultant, which I record.
The content of the interview is set by an agenda that we choose at the end of the preceding class.

- 10:10-10:15   Review class agenda set by preceding class
- 10:15-10:55   Interview language consultant [expanded below]
- 10:55-11:00   Set agenda for next time
- Before leaving: Confirm the session's transcribers

1.1.1   Interview of Language Consultant

While we interview the language consultant, the following will also be taking place:

- I keep track of data on the blackboard
- Students in class keep track of data in their notebooks
  (Buy a notebook specifically for this class)

It's very important for everyone to keep track of the data in their notebooks, especially those who will be transcribing the recordings. If anything is unclear at all on the recordings, your notes can be a vital reference!

1.2   Post-Class Period

Immediately following the end of the class, the following will take place:

- I will divide the recording of the class session into four approximately equal portions

- I will post each of those portions on the class language page
  (http://people.umass.edu/scable/LING404-SP16/Kalmyk/Recordings&Transcriptions/)

- I will send out an e-mail to everyone in the class letting them know the recordings are available. This e-mail will also notify students of which section of the recording they've been assigned to transcribe.

1.3 Transcription

- The appointed transcribers will download their assigned sound file from the website
- They will then transcribe the sound file, using the *XML template* available on the class language page (described below)
- Following their transcription, they update the class dictionary, by adding any new vocabulary and editing any older entries that need it (described below)
- Transcribers then e-mail their transcription to me, by the end of the following day.
- I post the transcriptions to the language page, alongside the original sound files.

---

2. Transcription, in Detail

In a class of this sort, it's best to set some rules/conventions for the transcription of our class interview sessions.

- Consequently, I'll be requiring you to transcribe those interviews using an *XML* template

---

A Brief Aside about XML

What is XML?
Basically, XML is a system that let's you *custom design* your own preferred way to represent some particular set of data.

Why are we using XML?
XML does two things for us:

(i)   It lets me set up and enforce some fairly **rigid guidelines for the representation of the language data** in our transcriptions.

(ii)   It's **open source**, which means *nobody owns it*. This means that, by representing our class data in this fashion, our data has a good shot of 'lasting the test of time'.

But, more about XML later… First let's go step-by-step through the transcription process…

The Three Major Steps to the Transcription Process:

- Download the sound file (Section 2.1)

- Open the sound file in Praat or Audacity (Section 2.2)

- Fill out the XML template as you listen (Section 2.3)

2.1   Downloading the Sound File

The first thing you need to do is download the file you've been assigned from the language page.

- Go to the class website (http://people.umass.edu/scable/LING404-SP16)
- Click on "Kalmyk Language Page"
- Click on "Recordings and Transcriptions of Class Interviews"
- Scroll to the date of the class session
- Click on the portion you are assigned to transcribe
- Download the file to your computer (click the icon at the top of the screen)

2.2   Open the Sound File

You can open and listen to the sound file in whatever program you like (*e.g.* iTunes). However, for the purposes of transcribing, it's helpful to view the sound file in either *Praat* or *Audacity*.

2.2.1   Praat

Praat is free software, designed by linguists, which can do all sorts of linguistically interesting things (create spectrograms, create synthesized speech sounds, do stochastic OT learning, *etc.*). You can download it here (a link also appears on the course language page, under "Class Resources"), and it works with both PC and Mac: http://www.fon.hum.uva.nl/praat/

Once you have Praat downloaded, open it. Now, do the following:

- On the "Praat Objects" screen, click "Open"
- Select "Read from file"
- Select the sound file; the title of the sound file should now appear in the "objects list"
- Click on the name of the sound file in the "Objects" list
- Click "View & Edit"
- A window pops up with a waveform and spectral analysis of the sound file (or rather, the part of the sound file that appears in the window)
- You can play the entire 'visible' selection by pressing the button marked 'visible part'
- You can also highlight a particular sub-portion, and play just that part by pressing the button that appears below it (refer to Praat help menu for more details)
- To replay a selection of the file over-and-over, just keep hitting the button!
- The buttons "in" and "out" let you zoom in and out to particular parts of the file.

2.2.2   Audacity

Audacity is a free sound-editing program. It can't do all the cool things that Praat does, but it provides a handy interface for examining the sound file. You can download it here (a link also appears on the course language page, under "Class Resources"), and it works with both PC and Mac: http://audacity.sourceforge.net/

Once you have Audacity downloaded, open it. Now, do the following (on a Mac):

- Under File, select 'Open'
- Select your sound file
- You will now see a simple waveform display of the file
- You can play the file by hitting the play button at the top of the window
- You can also highlight a particular sub-portion, and play just that part by pressing the play button.
- To replay a selection of the file over-and-over, just keep hitting the play button!
- To zoom in/out to subparts of the file, simply click View>Zoom-In or View>Zoom-Out

2.3   Filling Out the XML Template

As mentioned above, we will be sticking to certain set conventions regarding the content and form of our transcriptions of the class interview sessions. These conventions are encoded in the 'XML template' that we will be using throughout the class to write up our transcriptions.

First, you will need to download the XML template:

- Go to the class website (http://people.umass.edu/scable/LING404-SP16)
- Click on "Kalmyk Language Page"
- Click on "Class Resources"
- Under "Resources for Classwork", click on "XML Template for Transcription…"

Once you have the XML template, open it using either Notepad (PC) or TextEdit (Mac).

The template itself contains the instructions for how it is to be filled out.

2.3.1   Transcribing in IPA

As noted on the template itself, one of the fields that you must fill in is a representation of the utterance in IPA. This means that you'll need to configure things so that you can type in IPA. Here's how you do that:

Download a Unicode IPA Font:
- On the course language page, click under "Class Resources"
- Under "Resources for Transcription in IPA", click on either "Unicode IPA Font (Doulos)" or "Unicode IPA Font (Charis)" [I recommend the latter.]
- Instructions for downloading and installation appear there.

Download an IPA Keyboard Layout:
- Under "Resources for Transcription in IPA", click on "IPA Keyboard Layout (for Windows)" or "IPA Keyboard Layout (for Mac)"
- Instructions for downloading and installation appear there.

3.   Dictionary Editing

After you have transcribed the sound file using the XML template, and have sent the transcription to me, it's time to do your part in editing the class dictionary.

- For now, the class dictionary is a shared Google doc.

    o **I will share the Google doc with your UMass email address**

    o **Please register for Apps at UMass Amherst, if you haven't already (https://www.it.umass.edu/google-apps)**

    o At the end of term, I will convert a finalized version of the dictionary to PDF and post it to the course language page.

- Log into Apps at UMass Amherst and then access the dictionary in Google Drive or Google Docs

- Add entries for any new vocabulary encountered *in your part* of the session
  (The format for entries is explained on the dictionary document itself)

- If information gained during your part of the session indicates that older, existing dictionary entries need to be edited/updated, do this as well.

> **That's it! There's a lot of work at the very beginning – downloading and getting used to new software – but once that's done, things settle into a rather predictable pattern..**