Shannon Li

Project Plan

1. What data source will you use? #observations? time frame or other indications of data range covered

I will use the metadata of Nagoya University Conversation Corpus(https://mmsrv.ninjal.ac.jp/nucc/en/index.html), which is a corpus of Japanese natural conversation whose data are mainly recorded at around 2001.

I want to quantify and visualize the background information regarding the 198 conversation participants, such as their gender, age, birthplace and living area.

I also want to digitalize the metadata for the 129 conversation transcription documents. I want to quantify the scenario in which the conversation is taken place(private or public), the social positions of the participants as well as the relationships between the participants.

1. What will the visualization convey?

This visualization will convey the background information of each participant, each conversation, and the relationship between each participant of a conversation. By visualizing those material, my aim is to make it clear at first glimpse about how one person's speech style could relate to that person's background. In my follow-up research, my aim is to use this visualization to better understand how a natural Japanese conversation progress.

1. What two methods are you combining in your project?

I want to use Network analysis plot in displaying the participants and their relationships. It will be designed as directional plots to show how the speaker views the listeners' relationship comparing to their own. For example, older or younger, in a higher social position or not, being in close relationship or not. As these factors would impact the tone that the speaker chose to use.

Also, I will use geospatial plots to displays the birthplace and living area of the participants, as it will gives a hint about the dialects of that participants. In this plot, I will make sure that the age of each participant could be viewed in combined with the geographical background of that participants so that the language style of that participants could be deducted.

The plots I made in this project will be viewed in combination with the project of my digital text I course, in which I am digitalizing the corpus from plain text file into a digital system. I am planning to make the plots of this project interactive with that system so that in further research, the analysis process could be smoother with a clear visualization of the corpus.

1. Do you have any concerns for your planned project?

My main concern is on whether it is applicable to use a flow map to show a person's birthplace and the city that person has lived. Because even though the background information of some participants briefly provides these information, it does not cover all participants, which might impact whether it could honestly reflect on the possible dialectal influence on a person's language performance.