

Language and Society: Assignment 3

Zubair Abid, 20171076

Problem Statement:

An investigation of switching code and changing social identities in face-to-face interactions.

This study investigates the code-switching and identity construction among a (filtered) trilingual population in Kolkata reasonably conversant in English, Bengali, and Hindi (“reasonably conversant” does not imply complete fluency in any of the languages).

It seeks to observe, and to establish if there exists any change in behaviour and language used across age groups in their interactions with three perceived classes of society; namely, a “Higher Class” - persons with outwards indications of affluence and considered respectable and influential among society, a “Middle Class” - the ‘general class’ of the average person on the road, and a “Lower Class” - those exhibiting financial, mental, physical poverty and looked down upon by most in society.

It also observes if there is any code switching if the surveyor seems to be uncomfortable using the assumed language projected onto them.

A filtered subsection of trilingual speakers is taken, namely families (children, parents, and grandparents) of school-going children where the primary medium of instruction in the school is English. This group is taken on the assumption that the large majority of those surveyed will fall in the “Middle Class” category themselves (as we’re limiting our study to said group) and any outliers are statistically insignificant.

Data Collection:

Data is collected by survey; people dressed up as typical members of the three defined social classes strike up a conversation with somebody, asking similar questions (such as “What’s the time?”, “Where’s the nearest station?”, and the like) and noting down their responses, first language used, and behaviour.

Since we’re investigating the code-switching at play, we need a ‘neutral’ first question such that the response can be in any of the three languages.

Informants for data collection:

A filtered subsection of trilingual speakers conversant in English, Bengali, and Hindi is taken, namely families (children, parents, and grandparents) of school-going children where the primary medium of instruction in the school is English.

This must be done very carefully to ensure the specific combination of trilingually competent middle class families with children with English as their primary mode of instruction. Since it is not practically feasible to accurately verify such credentials via question/answer or judging by sight, we choose to collect data from areas where such an intersection of demographics is known to occur, namely specific classes of schools where the primary medium of instruction is English.

It is not necessary that all persons should be grammatically fluent in all of the languages, and the languages need not be limited by ordering in terms of relative fluency - one may be more comfortable in Bengali than the others while somebody else may find Hindi most comfortable.

The informants are divided by age groups: 6 - 20, 21 - 35, 36 - 50, 51 - 65, and 65+. Age classification is done by visual estimation,

Sample Size:

About hundred people from each age group are selected for analysis against one another. The reason hundred are chosen is a matter of compromise; we need more data points for accurate analysis, but surveying a large number of people becomes practically infeasible. Hundred is large enough number for basic correlation, with scope for future secondary research if required.

Data Collection Method:

We collect data by asking people a variety of questions, and noting down:

1. First language used to respond
2. Behaviour while responding (polite, curt, etc)

The first question is asked as language-neutrally as possible, to elicit a response based on what the informant assumes is the questioner's preferred language. This observes if there is any change in language used in response to the perceived class.

After that, the conversation continues in the given language where the questioner pretends to be uncomfortable using the language, and we observe if the respondent switches to any other language or not, and which language they switch to.

In order to get the data for all three classes, each person is approached (with reasonable intervals of time) by three surveyors representing the different classes, all of whom strike up conversations. After every conversation, the behavior and language usage is noted down, and once all is done, compared against one another.

Material Preparation for given collection method:

Question selection is a crucial portion of the study -

1. It must be such as to reduce Observer's Paradox as much as is possible
2. The first question should be selected such that the informant can respond in any language they wish to.

Since the informant here would not know that they're part of a study (this is potentially ethically ambiguous, but not really. See: Labov's New York Study), care should be taken to ensure that no suspicion is raised, conversation should be as normal and free flowing as possible.

Medium of method in Data Collection:

Data Collection is done verbally via survey and later noting down what was observed. Each person is approached (separated by reasonable intervals of time) by three surveyors representing the different classes of society being considered here, and their behaviour/language usage is noted down by each. Each surveyor must also keep track of which person is being studied, and ensure that everybody surveyed interacts with all 3 classes of society.

Data Analysis:

The first analysis done is a comparison of first language use for initial communication. For each age group, we count the number of people who begin in each of the three languages for every given class of people surveying. This information can be plotted on a bar graph taking the y axis as the number of respondents and the x axis as divided into kind of responses per given class. A similar approach is applied for the remaining data as well.

A Sample visualisation is given below (the data is filler and not representative of any actual data that is yet to be collected).

Number of responses by language and age

