

Language and Society

Individual Analysis Report for project

Social Stratification

Name – Suyash Vardhan Mathur

Roll no – 2019114006

Overall Understanding of the Project

While our project initially just mentioned Social Stratification, the approach that we have taken to understand it is to observe characters and situations from Movies+TV Shows+Interviews+Advertisements, and we have made a certain set of Hypothesis regarding Social Stratification, based upon dimensions like:

- Wealth/income
- Age
- Gender
- Religion
- Region/Language
- Occupation

Now, in order to evaluate the correctness of this hypothesis, we gathered data from real world if people of a particular background actually show Stratification as per our initial hypothesis. However, since we cannot ask people to identify the features involved in our hypothesis, we indirectly asked them the same by asking their relatability to advertisements, which exhibited those characteristics. We work under the assumption that people relate to those advertisements which involve the same features that their Stratified category relates to/uses. Thus, based on the responses from the form, we determined the accuracy and correctness/falseness of our hypothesis.

In order to form the hypothesis, we analysed 50 ads on selected socioeconomical dimensions and linguistic features(along with some non-linguistic features). Next, we conducted a survey where we collected a set of ads based on certain socio-stratification elements visible in them, that *how similar was the language they speak and the language in the given advertisement?* We analysed these collected responses, and were able to make some conclusions based on this analysis.

Major Contributions

- Made the form for collection of advertisements[along with others in the team] that asked the top 5 advertisements that people could relate to.
- Analysed the form [responses](#) with the team, sanitised them with the team, and found common elements in the form which could be related to social stratification, like how students felt relatability to Udemy ads, and wrote analysis for Kia ads.
- **Read 2 research papers** - [Dialect Differences and Social Stratification in a North Indian Village](#) and [REVISITING KHALAPUR i : LANGUAGE VARIATION AND SOCIAL STRATIFICATION 50 YEARS LATER](#), and analysed them to find the Linguistic parameters we can further look at while working on our project. The parameters in the paper that helped in the project were:
 - *Village and Urban speech had vast differences among them.*
 - *Castes expressed differences among themselves with the help of language/phonetic features, such as Chamars and Rajputs had vast differences in speech.*
 - *Phonetic features were used for social stratification, like many chamars used /ə/ in forms like /batau/ instead of /bUtau/, untouchables of all three castes having nasalized /l/, /ä/, /u/, nasalisation of oral vowels by lower castes, which was considered as an indication of their caste, chamar and sweeper speech being different in terms of contrast between /a/ and /ə/, /ul/ and /u/.*
 - *Vocabulary differences were also used as an expression of stratification, like in case of Hindu-Muslims.*
 - *Other specific phonetic features, like diphthongs /ai/, /ui/, /oi/ before consonants, nasal retroflexion, nasalisation, breaking of consonant clusters proved to be useful pointers for looking at phonetic features for our analysis*
- Overall, I analysed 11 advertisements(17 different characters), with detailed analysis of 3 ads(Kia, Dhanya Seeds, Rin(Upper Class)) (in analysis doc) and other ads in the Data & Analysis of Advertisements - Dimensions.csv and Features.csv
- Also analysed various characters from **Mirzapur TV Series** and formed conclusions based upon clips which contrasted sub-urban/village populace(Given in detailed analysis doc).

I observed different kinds of Codemixing that was very apparent among people of different Upper/Middle class setting, regional accent differences, specially in case of Rural Uttar Pradesh, like using of lexical features like *bhaiya ji, kahe*, stretching u like *bahut* -> *bahuut*, phonetic features like converting /j/ -> /s/ and z -> ďz, regional slangs, like those local to UP like *bhaukal, tafri*. I also

observed contrast in the English words spoken by Rural/Urban, like rural characters speaking congratulatory words for congratulations, option for 'option', using pure Hindi words like *कृषि* instead of *खेती*, using *कृतज्ञ*.

- Decided the advertisements for the final survey form along with the team, and how the form would be presented. We all picked up 1-2 advertisements to include in the form, and to be presented.
- Did the quantitative analysis for the **Forevermark** and **Airtel** advertisement, and drew conclusions with respect to the Rural/Urban Distribution, Range of Income, Gender which were visible in Forevermark, and with respect to the Age targeting (teenagers), wrt Mother Tongue in case of Airtel ad. [Given in detailed analysis doc]

Challenges

- The data that was collected was heavily biased towards people living in the cities, and Telugu and Hindi speakers comprised a lot more among the people who filled the form than people of other language.
- We felt that there was some level of misunderstanding of the questions – like taking the narrator's voice into account for features as well apart from just the features of the speech of the characters involved in the advertisement.
- The accuracy with which people looked at the Linguistic features was rather loose, and people weren't able to accurately identify the accents which were targeted towards the rural populace.