Aashwin Vaish
2019114014

Language and Society
Project
<u>Social Stratification</u>
Individual Report

Overview

The project is for understanding social stratification through popular audio-visual media including advertisements, TV shows, Movies etc.

The project requires us to form a Hypothesis using thorough analysis of language in TV shows and movies, and try and confirm it through a survey form, which asks for person's social factors that we think relevant, and the ads they commonly relate to.

We analysed around 50 ads on selected socioeconomical dimensions and linguistic features (and a few non-linguistic features to supplement). We then conducted a survey where we asked people for some socioeconomic information and gave them some ads which we decided based on the analysis we did, and asked them to rate *how similar they thought the language they speak and the one used in advertisement were similar?*. We analysed the data and keeping the biases in mind were able to draw some light conclusions.

Contribution / Analysis / Observations

When survey form #1 was conducted, a lot of *troll responses*, *misfilled responses*, etc had to be filtered out and the data had to be sanitized which was done in part by me.

In total I was able to analyse 13 advertisements collectively in Hindi and English, 1 Interview (3 minute interaction scene) from a Hindi comedy-talk show, including a transcript (presented in earlier submissions) of the conversation.

Sample Analysis *[Rest are available in the Data Analysis.csv file in GitHub Repo marked with their respective analyzers]*

<u>Airtel (Har ek friend zaroori hota hai)</u>
- /r/ -> /ɹ/ [in english environment]
- Heavy English code-mixing
- Youth Slangs, Mild vulgarity
- Mumbai Accent
- College and similar
- Modern
- The ad is a song

Aashwin Vaish
2019114014

I was chiefly able to observe code-mixing / switching , regional accent differences, phonetic / lexical features. Frequency of code-mixing increases with decrease in Age, increase in Wealth, when person lives in a City, or how different Mother tongue is to the lingua franca. Frequency of code-switching followed much the same trend, but age didn't seem to have a role here.

I collaborated on the presentation and design of the second survey form to come up with ways to frame the questions and display advertisments to respondents. Then we all had to pick 1 ad from the IL we speak (except for Malayalam, where Jerrin had to pick all 3) so that we can ask the respondents the similarity question (as mentioned previously).

Apart from this we all had to analyse the responses for one or two ads of the language we had chosen, and I was aggregating responses for Blackberry and Pediasure Ads from English and Hindi sections respectively (Form Responses, which are uploaded on GitHub as well).

Challenges

- There is definitely a bias from the advertisers towards caring more for linguistic accuracy when dealing with "non-standard" Hindi, or targetting a minority group (e.g. Rural groups, Ethnic folk etc.)
- South India has a common culture where a lot of people from villages get their education in cities. I being one of the form designers, and being North Indian, was not aware of this in time so we weren't able to design the "Region - City/Town/Village" question to accommodate this.
- Telugu speakers formed the majority of responses. It seems we weren't able to make clear in the form how much proficiency a person should have in order to say that they "knew" the language. A lot of people claimed to know a variety of Hindi that I really don't feel like they would (e.g. many urban telugu people could "Highly Relate" to Rural UP/Bihar Hindi).
- People in general are not vigilant enough to notice small phonetic changes, and even when presented in a string of words out of which some are not native to their dialect (slangs etc), they still claim to have understood the utterance (unless slang is the focus), which distorts the results.