Report about the previous stuff that we were doing.

* Interviewee and interviewer switching pattern
* Why both of them follow the same sort of pattern in their word choice and language choice behavior? Is that something because of the influence of the interviewer over interviewee?
* Word distribution for the switch point
* Word distribution after the switch point
* You can do the previous two bullets for unigram, bigram, trigrams.
* One of the phenomenon that can be checked out is interpretation and balance of the switch language, isn’t it interesting? For example, there is the same number of switches from Swahili to English and English to Swahili. What can tell us this symmetry?
* Running a new code switch prediction algorithm based on the most frequent bigram, trigram, and unigram that we see in our dataset. In other words, the word itself can be used as a feature now.
* Comparing interviewer and interviewee plots and finding interesting pattern by comparison.
* Making the plots nicer and better to read.
* Corrections to the bar plots of the first presentations and removing the lines that connects the bars.
* Plotting the word distribution for immediate switches. For example the words that happen to be after a switch point and they themselves are also switch points. Like unigrams, bigrams, and trigrams.
* Analyzing the results and making the interpretation is something very important to us.
* Swahili annotation using crowd sourcing.
* Frequency graph of each utterance length is something that we might need it to have. How many sentence we have that have the length of 1 in Swahili, 2, 3, etc. The same for English.
* Is it true that one of the interviewers imposes on the interviewee the style or choice of the words? How can we check this in our interview data and how can we validate this with the jamiiforum data?
* Enforcing selection of the switch words, is it somehow enforced by the interviewer to the interviewee to choose some specific words for the switch or not? You can check this with the interview data and validate through the JamiForum data.
* Put examples of switching or words in each plot that you want to describe.
* Same analysis on JamiiForum and validating our results.
* Is Jamii forums follows the same switching pattern as the interview file or are they using the same kind of words for switch points are the questions that we can ask?
* How can we use the social meaning annotation from our interview data?
* In the plots move the ones that are mirror of each other next to eachother and study if they have similar numbers and if “yes” why ? for example sw>en and en>sw.