Devina Dhawan  
Chantiny Griffith  
Report – NLP Project 2

For word order, we came up with rules to account for local ordering of words, such as a determinant not having a noun or adjective to follow it, or a noun followed by an adjective. We selected the most common mistakes and took off points for each occurrence. Later we realized we could have used a pre-written program to find word dependencies, but the way we programmed worked well and we were able to learn more about the precise grammar rules for which we were testing.

The main issue that we had with subject-verb agreement was highly technical. We tried to associate every word with its corresponding tense, and its part of speech. This may have over complicated the problem and caused a lot of issues for us. So we more or less hard coded a lot of what would be a proper subject verb agreement for any given sentence, and then awarded or took away points based on if it was good or bad sentence.

As for the verb tenses, we realized that there were certain words that are auxillary, and so we used what we knew in order to iterate through the all the words in the paragraph to figure out if they used the correct verb tense for every word they used. We were able to learn that a lot of the time the student understood their task and were able to use the correct tenses but then they would have other errors which would give them incorrect scores.

We had some trouble implementing our coherence checks. Ultimately, we decided to count the number of third person pronouns that are used, check that there is a possible antecedent, and check that the gender of the pronouns match. We realize that gender can change if someone is talking about multiple people, but when we attempted to implement this it flagged more things incorrect than it should have, taking away points from correctly formed sentences and not from incorrectly formed sentences. Since the essays were short, this solution mostly worked. In a larger implementation, more rules would have to be added to account for subject shifting.

In order to grade if a paragraph was on topic or not, we basically needed to understand what words one would hit if they were discussing their family or their background. Inorder to accomplish this, we had to read through a lot of the papers to understand clearly what the requirement was. In larger implementation we would need to have a larger array of words, preferably given to us through a dictionary source.

The grading we designed for checking criteria 1D, sentence formation, was difficult at first. We had already implemented 1A and had to spend some time discovering how the scores overlapped and how to not deduct points for identical problems, but still give accurate scores. We implemented a call to the parser and used the output of that to determine if there were sentence fragments or poorly formed sentences. This criteria was about overall sentence structure more than the correctness of ordering between two individual words, so we deducted points for each problematic fragment.