

**TO:** Philipp Koehn, Adi Renduchintala  
**FROM:** Sarika Halarnakar, Sindhuula Selvaraju  
**DATE:** March 25<sup>th</sup>, 2016  
**SUBJECT:** Final Project Proposal

---

### **Introduction**

For our final project we will be creating a machine translation homework assignment on reordering. Languages differ in many ways based on word order alone. Word order defines relationships between words. Some classifications can be made by naming the typical order of subject, verb, and object, but there are also many other differences in word orders such as Adjective/Noun placement, Noun/Relative Clause placement, and in some languages the word order is not fixed. In machine translation, it is important to preserve word order because the meaning of a sentence can change depending on word order. For example, the difference between “I had my house cleaned” and “I had cleaned my house” is an important distinction to make as the word ordering changes the meaning of the sentence.

We will be assembling a basic implementation that provides baseline metrics. Students will need to improve these metrics in their own implementations. Similar to our previous homework assignments, we will provide some implementation suggestions to help improvement.

### **Project Outline**

We still need to read more papers to solidify what models and algorithms we will implement, but we are currently planning to provide a basic grammar ruleset that student would have to add to in order to build a fully comprehensive ruleset and improve their results. We plan to implement our project using python.

### **Relevant Papers**

1. [Pre-ordering of phrase-based machine translation input in translation workflow](#)
2. [A Word Reordering Model for Improved Machine Translation](#)
3. [Automatically Learning Source-side Reordering Rules for Large Scale Machine Translation](#)

### **Experimental Design**

**Input:** hyp1 ||| hyp2 ||| ref : hyp1 and hyp2 in English and ref in French

**Output:** 1 or 2 depending on which sentence is more correct.

**User Scoring Rubric:** We will build a compare\_results program, similar to the one provided to us for our other homeworks. The user will use this program with their output to check if their accuracy is better than that of the program given by us.

**Evaluation Measures:** Compare with predetermined result file which also contains values 1 or 2 corresponding to each line of input.