Mariana Graterol Fuenmayor

**LING83600**
**HW2 - GRAPHEME TO PHONEME CONVERSION**

Overall, this assignment was pretty straightforward. My main challenge with the main goal was the fact that the official documentation on `fairseq` is not the most comprehensive. In order to find the full selection, I resorted to looking at their implementation of the command-line tools directly in the github repository, and, finally taking the output of running `fairseq-train -h` directly in the terminal as my main source of information. This proved to be a little challenging, because I was picking arguments just by their name instead of fully understanding what their purpose was in the training process.

To understand the predictions document output by `fairseq-generate`, I mainly followed this resource. Following the explanation given there, I chose to use the line prefixed by `D-` as my prediction. For our assignment, since we do not use byte pair encoding, the tokenized hypothesis `H-` and the detokenized one `D-` should have been the same, and I did check to make sure it was the case, but using the latter seems to make the script more generalizable.

The lack of documentation also proved to be a challenge when I tried using the transformer model with the suggested paper's parameters to try the stretch goal. Exploring with different arguments, my predictions went from being /a/ to /a a a a a a/ to something that made sense. After a few tweaks, I finally achieved a WER of 41, significantly worse than my previous WERs of 16 and 21. This could be due to the fact that I maintained the number of updates at 800, since it seems that the transformer gets better later on, or perhaps I might still be missing something in the arguments. Maybe all I need is to do more reading about how these neural models are implemented in the backend.