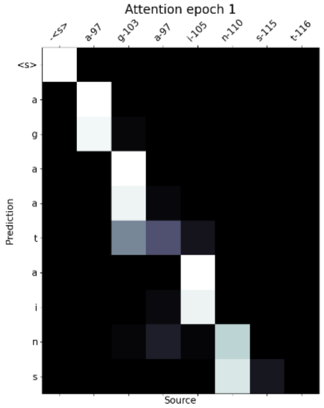
Refael Shaked Greenfeld 305030868

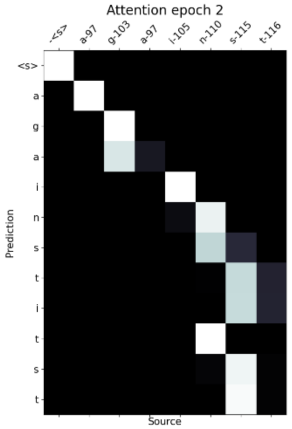
Part 3:

Attention Plots for the word against Without teacher (greedy):

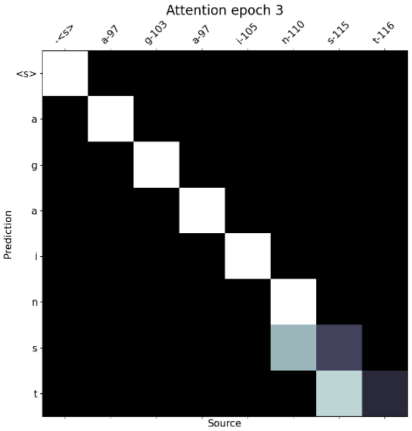
Attention epoch 1



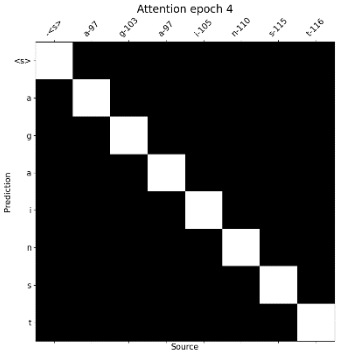
Attention epoch 2



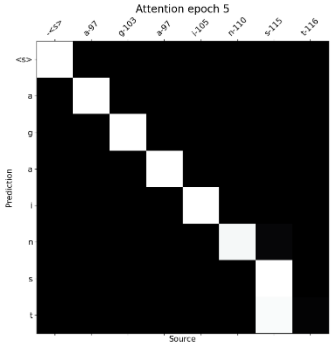
Attention epoch 3



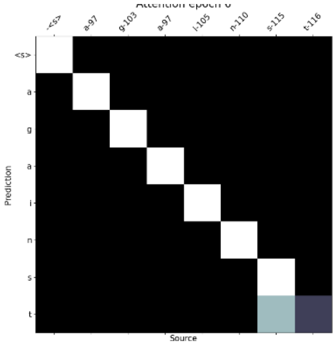
Attention epoch 4



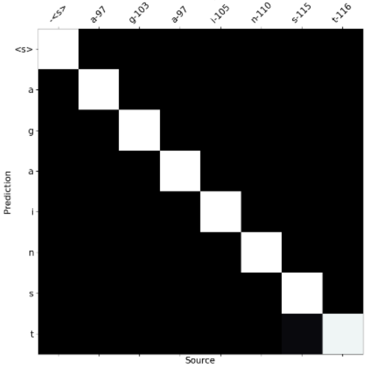
Attention epoch 5



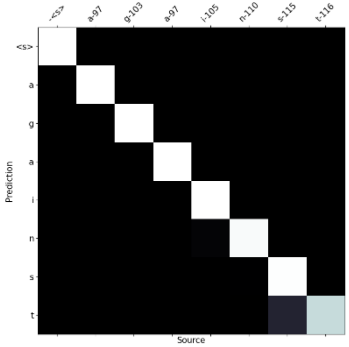
Attention epoch 6



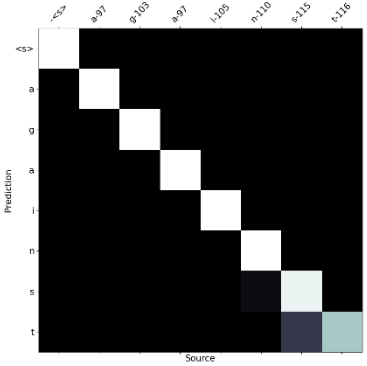
Attention epoch 7



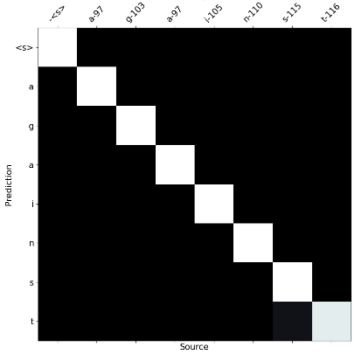
Attention epoch 8



Attention epoch 9

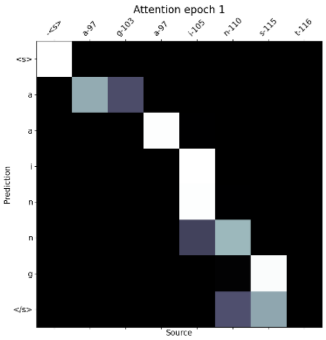


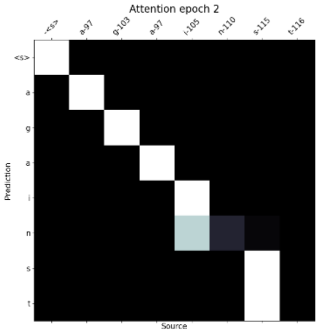
Attention epoch 10

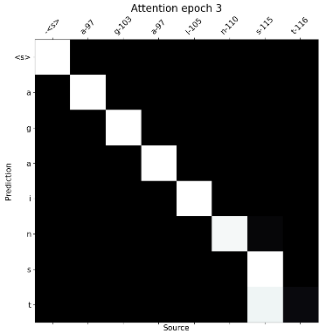


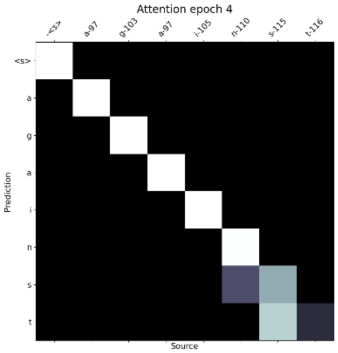
Refael Shaked Greenfeld 305030868

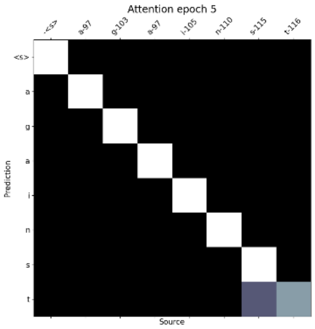
Attention Plots for the word a g a i n s t teacher:

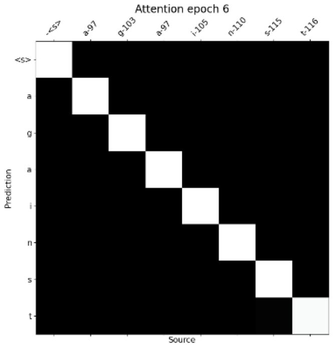


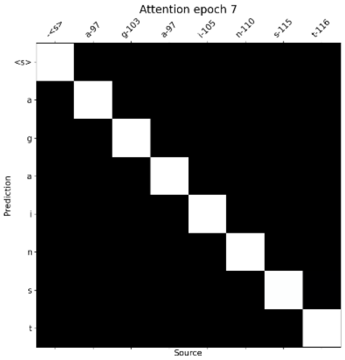


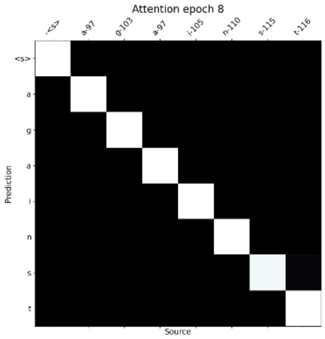


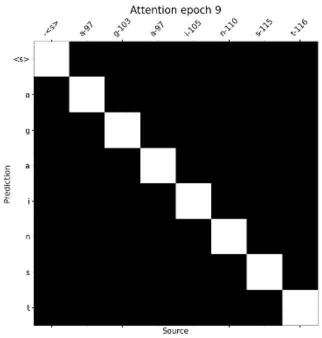


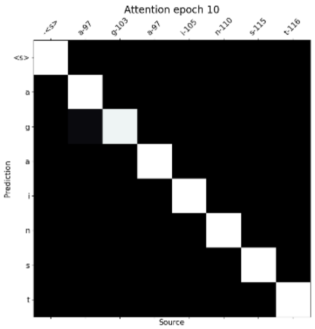












Refael Shaked Greenfeld 305030868

General Question:

First, I plotted 2 set 1of 10 graphs, one as in the training, with the “teacher” with fixed length, and the other like in inference with a changing length. I wasn’t sure what you meant, I submitted the one with the teacher, hope its ok (they are pretty similar and the idea is the same).

For the question,

The attention weights are changing by every epoch and becoming more and more diagonal (weight is where the weights are higher) which makes sense because the weights are improving and the input and output are aligned so this what we would expect. In the last epochs we can see a diagonal white line which means that evet index effect the most on the respective letter. Interesting to see how the mixture of effect is in the ‘t’ and ‘s’ in the end, we can explain it by the fact that many words ends with ‘t’ and ‘s’ so the attention weights are able to learn this correlation in the appearance in the end of a word. And the general trend is diagonalization in the attention weights which means it is able to learn the correct “soft alignment”.