

**Information Bottleneck** - the final hidden state of the RNN has to capture the meaning of the whole sequence, including the first word

- If the first word is important for decoding the last word, the information has to go through many transformations to get there
- Some of the information gets lost over time

