# Subword Models

## Shubham Gupta

## February 27, 2020

# 1 Phonetics and phonology

- Phonetics is the sounds steam

- *Phonemes*: Small sets of units used commonly by humans to communicate with each other

- *Morphenes* Smallest sematic unit. Helps break down words into smaller units that have distinct meanings

- Difficult to work with morphenes though. Hence use character n-grams.

- Need to handle large, open vocab for word based models

-