Am I Stressed? Detecting Stress in Bengali
Shohini Bhattasali
Cornell University

Objectives
- Train a model to detect stress in Bengali, using data from a Bengali speech corpus
- Compare my results against existing accounts of Bengali stress

Existing accounts
- Proposal 1 states that stress falls on the initial syllable (Hayes & Lahiri 1991, Chatterji 1921, Ferguson & Chowdhury 1960).
  (1) ˈbəri ‘house’
  (2) ɖuʃeːilo ‘entered’
  (3) ˈmaːkəliɛn ‘you had mashed’
- Proposal 2 states that stress falls on the initial syllable only if it is heavy (CVC), otherwise on the second syllable (Shaw 1984).
  (4) ˈsɔŋʃar ‘family’
  (5) aˈkaʃ ‘sky’
  (6) koˈbit ̪a ‘poem’
- Both proposals agree that stress may fall on the first syllable but only Proposal 2 predicts that stress may fall on the second syllable.
  (7) ˈbatʃa ‘type of candy’
  (8) baˈtʃa ‘type of candy’

Proposal
- Test the two existing formal accounts of stress by training an ASR (Automatic Speech Recognition) system to detect stress cues in a Bengali speech corpus, Shruti (Mandal et al. 2011).

Experiment
- Using Kaldi (an ASR toolkit created by Povey et al. 2011), the sound waves from the corpus were converted to their corresponding vector representations.
- Through OpenFST (a weighted finite-state toolkit created by Allauzen et al. 2007), two finite state models were trained on a phonetic dictionary to map these vector representations to the transcript of an utterance.

Model 1:
- Stage one: Trained on a phonetic dictionary with stress marking on the first syllable
- Stage two: Trained on a phonetic dictionary which had three variants for each word: no stress marking, stress marking on the first syllable, and stress marking on the second syllable.

Model 2:
- Stage one: Trained on a phonetic dictionary with stress marking on the second syllable
- Stage two: Trained on the phonetic dictionary which had three variants for each word: no stress marking, stress marking on the first syllable, and stress marking on the second syllable.

Results
- Although initially trained on different phonetic dictionaries, in stage two both models detected stress on both the first and second syllables.
- The percentages are given in the bar graph:

Conclusion
- Model 1 & Model 2 detect stress cues on both the first and second syllable and thus, this suggests word-level stress is not strictly word-initial (supports Proposal 2).
- Finite-state models can be trained to detect prosodic and suprasegmental phonological features in natural language.

Selected References