Changes we need to or should make, somewhat prioritized

* instead of just removing punctuation, should keep it as a separate word, at least in the case of EOS markers
* add more features
  + all caps
  + initial cap
  + hashtag
  + in sentiment dict
  + POS tag

Tag/chunking notes

To get started in interpreter:

sys.path.extend(['/users/cindi/semeval'])

import semevalTask4

Try train and test on the trial data first to see how it works.

ChunkParse score on restaurants-trial 80/20 split:

IOB Accuracy: 89.0%

Precision: 75.0%

Recall: 24.0%

F-Measure: 36.4%