

# Assignment 2

L545

Due Monday, February 4

1. Consider the following Japanese data:<sup>1</sup>

tabeta	'X ate Y'
aketa	'X opened Y'
tabesaseta	'X made Y eat Z'
akesaseta	'X made Y open Z'
taberareta	'X was eaten'
akerareta	'X was opened'
tabesaserareta	'X was made to eat Y'
akesaserareta	'X was made to open Y'
tabesasenai	'X doesn't/won't make Y eat Z'
tabenai	'X doesn't/won't eat Y'
tabesaserareru	'X is/will be made to eat Y'

- (a) Based on this data, give the Japanese morphemes for the following English translations:
- 'open'
  - 'eat'
  - passive (+PASS) marker ('... be VERB-ed', e.g., 'They were opened/eaten')
  - causative marker (+CAUS) ('... make X VERB', e.g., 'Robin makes Tracey laugh')
  - non-past marker (-PST) (present or future tense)
  - past marker (+PST)
  - negative marker (+NEG)
- (b) Provide (i.e., properly define) a finite-state transducer (FST) to recognize and analyze these words
2. Do question 3.2 on p. 81 from Jurafsky & Martin, extending the transducer to deal with *sh* and *ch*.
3. Draw a FST which accounts for the following spelling rules, where the input is a morphotactic form and the output is an actual English word:
- $y \rightarrow ie / \_ \hat{s} \#$
  - $y \rightarrow i / \_ \hat{ed} \#$
- Remember to account for words with non-final *y* (e.g., *you*, *dystopia*) and to disallow the strings *ys* or *yed* ending a word. You do not need to worry about interactions with other spelling rules.
4. Do question 3.5 on p. 81-82 of Jurafsky & Martin, involving the Soundex algorithm.

---

<sup>1</sup>Question adapted from *The Language Files*.