Lab 2 – Linguistics I

1. Learning Objectives

In this lab we are going to:

- Know more about linguistic structure, analysis and data

2. Formal Grammar

Define a formal grammar (grammar rules + lexicon) that can be used to analyse/generate the following sentence:

He had lunch at the restaurant.

3. Formal Grammar (recursion)

Extend the grammar so that it can be used to analyse/generate the following sentence:

He had lunch at the restaurant with the big windows on the corner of the street.

4. Part of Speech Tagging

Use the PoS tag set from the slides to annotate each word in the following sentence with the correct part of speech:

He had an expensive, but very good lunch at the Thai restaurant with the big windows that is opposite the church.

5. Part of Speech Tagging, Parallel Corpus

Translate the sentence of question 3. in a language of your choice and provide:

i) cross-lingual alignments between the words of the sentences
ii) part of speech tags for the words of the sentence in the other language (using again the PoS tag set from the slides)

6. Morphology

Stemming and lemmatization are two important morphological processes that are useful for many NLP applications.

In simple words, both processes can be used as pre-processing steps to transform a word into its base form but using different techniques.

Stemming is the process of heuristically removing the affixes of a word, to get its *stem* (root).

6.1. Can you find the stem the following words: amusing, amusement, and amused?

Lemmatization is the process of morphologically analysing a word that returns its *lemma*, which is the form in which a word appears in the dictionary.

6.2. Can you get the lemma for the following words: amusement, amusing, and amused?

Some languages have a complex morphology such as the Arabic language. Consider the word

وسيكتبونها

/wasayaktubuwnahA/

6.3. Can you guess the PoS tag for this word?
6.4. Provide the lemma & stem of a word in a language of your choice?
6.5. Name some applications that might use these morphological processes?

7. Phrase/Constituency Structure vs. Dependency Structure

Provide a phrase/constituency structure (tree) and dependency structure (tree) for the following sentence.

We ate at a cheap restaurant on the corner of the street.

8. Synonyms, Hyponyms, hypernyms

Use WordNet (http://wordnetweb.princeton.edu/perl/webwn) to provide synonyms, hyponyms and hypernyms for the following nouns: *cat, work, deal*

Please note that in the web interface you will need to click on 'S' in order to show hyponyms/hypernyms.