# Computational Linguistics Introduction

Computational Linguistics: Jordan Boyd-Graber & Philip Resnik

University of Maryland

Warmup

# Agenda

- (Re)introductions
- Breakout: Questions from you
- · Getting homework code checked out
- Warmup exercise

## Instructional Staff

- Philip
- Jordan
- Khanh
- ???

### **Breakout Session**

- Introduce yourselves
- What questions do you have
- Answer yourselves
- Otherwise, create a list
- Nominate a spokesperson

## Checkout the Homework Code

https://github.com/Pinafore/cl1-hw/tree/master/limerick

#### No NLTK

```
jbg-mini:cl1-hw jbg
jbg-mini:limerick jbg
ls
README.md limerick.py tests.py
jbg-mini:limerick jbg
python3 tests.py
Traceback (most recent call last):
   File "tests.py", line 2, in <module>
        from limerick import LimerickDetector
   File "/Users/jbg/repositories/cl1-hw/limerick/limerick.pumport nltk
ModuleNotFoundError: No module named 'nltk'
```

```
jbg-mini:limerick jbg$ pip3 install nltk
Collecting nltk
 Using cached nltk-3.5.zip (1.4 MB)
Requirement already satisfied: click in /usr/local/lib/pyt
Requirement already satisfied: tqdm in /usr/local/lib/pyth
Collecting joblib
 Downloading joblib-1.0.0-py3-none-any.whl (302 kB)
     Collecting regex
 Downloading regex-2020.11.13-cp37-cp37m-macosx_10_9_x86_
     âŰĹâŰĹâŰĹâŰĹâŰĹâŰĹâŰĹaŰĹaŰĹaŰĹaŰĹaŰĹaŰĹaŰĹaŰĹaŰĹaŰĹaŰĹaŰĹa
Building wheels for collected packages: nltk
 Building wheel for nltk (setup.py) ... done
 Created wheel for nltk: filename=nltk-3.5-py3-none-any.v
 Stored in directory: /Users/jbq/Library/Caches/pip/wheel
Successfully built nltk
```

## No data?

```
Resource cmudict not found.
Please use the NLTK Downloader to obtain the resource:
>>> import nltk
>>> nltk.download('cmudict')
```

#### Install data

```
jbg-mini:limerick jbg$ python3
Python 3.7.2 (default, Jan 13 2019, 12:50:01)
[Clang 10.0.0 (clang-1000.11.45.5)] on darwin
Type "help", "copyright", "credits" or "license" for more
>>> import nltk
>>> nltk.download('cmudict')
[nltk_data] Downloading package cmudict to /Users/jbg/nltk[nltk_data] Unzipping corpora/cmudict.zip.
True
```

You'll need to do this later for Penn Treebank, WordNet, etc.

# **Expected Failure**

```
FAIL: test_syllables (__main__.TestSequenceFunctions)

Traceback (most recent call last):
  File "tests.py", line 30, in test_syllables
    self.assertEqual(self.ld.num_syllables("letter"), 2)

AssertionError: 1 != 2

Ran 3 tests in 1.706s

FAILED (failures=3)
```

### Thou Must

## Challenge

Find all examples of "thou \_\_\_\_t" in the bible; what are the most frequent?

- nltk.corpus.gutenberg
- import re
- FreqDist or Counter