**Building a Child Language Analyser**

Aniruddha Indurkar

**Index**

**Contents**

|  |  |  |
| --- | --- | --- |
| **Sr. No** | **Title** | **Page Number** |
| 1. | Introduction | 3 |
| 2. | Instruction | 4 |
| 3. | Limitations | 7 |
| 4. | Code support | 7 |
| 5. | References | 7 |

**List of Images**

|  |  |  |
| --- | --- | --- |
| **Sr. No** | **Images/Screenshots** | **Page number** |
| 1. | Output of Task-1 | 4 |
| 2. | Unsuccessful output of Task-1 | 5 |
| 3. | Output of Task-2 | 5 |
| 4. | Output of Task-3 (a) | 5 |
| 5. | Output of Task-3 (b) | 6 |
| 6. | Saved image from Task-3 | 6 |

**Introduction**

**Name:** Aniruddha Indurkar

**Python version used:**  Python 3.6

**Libraries used:**  os, re, numpy, pandas, matplotlib

**OS used:**  MacOS

**Windows/Linux Compatible**: **Yes**

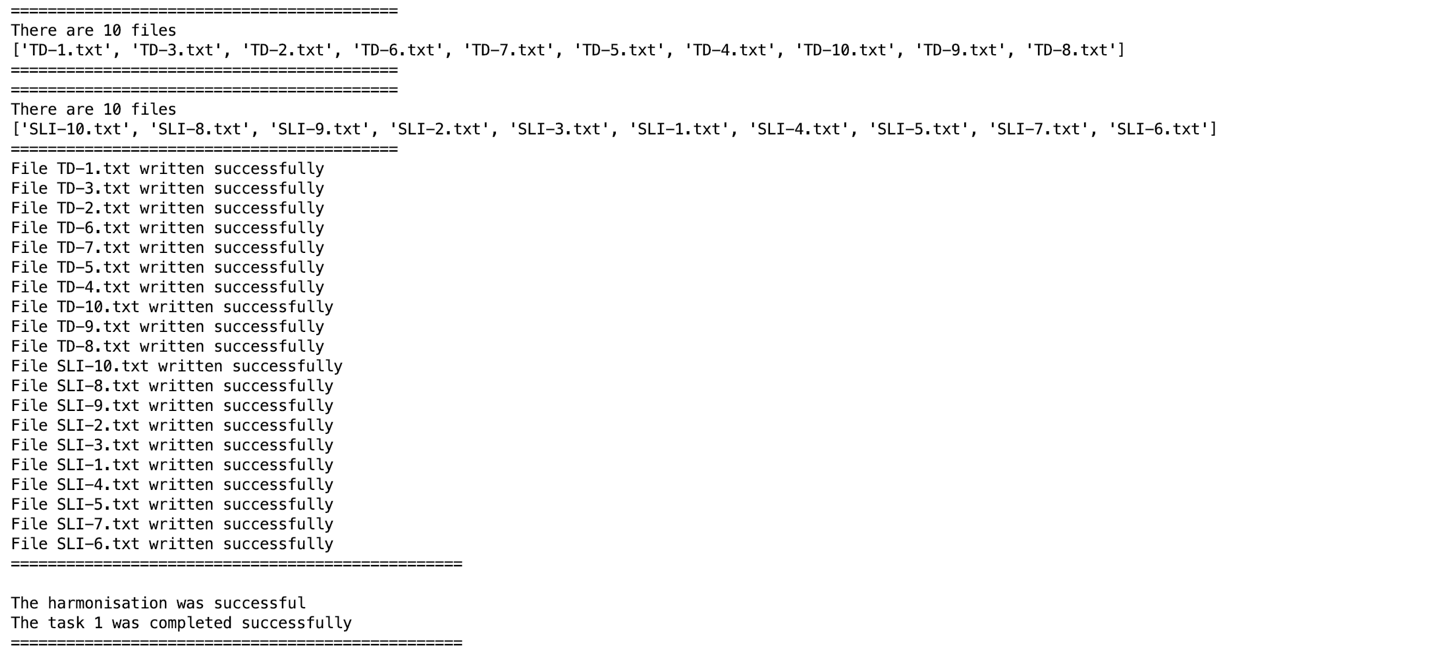
**Instructions**

1. We make use of the scripts in the terminal or the command prompt.
2. We need to download the data from moodle and keep ENNI Data set folder in the same folder as the working directory for python.
3. We assume that the folder name should be 'ENNI Dataset', otherwise, the task will not be successful.
4. Inside this folder, again, we assume that there are two folders named 'TD' and 'SLI' and they contain files in the format *'TD-XX.txt'* and *'SLI-XX.txt'*. In total, there should be 20 files, ten files in TD and ten in SLI folder in-order for the harmonisation to be successful.
5. We make sure if the numpy, pandas and matplotlib are installed. If not please run them using the below commands in the terminal:

a. pip install pandas

b. pip install numpy

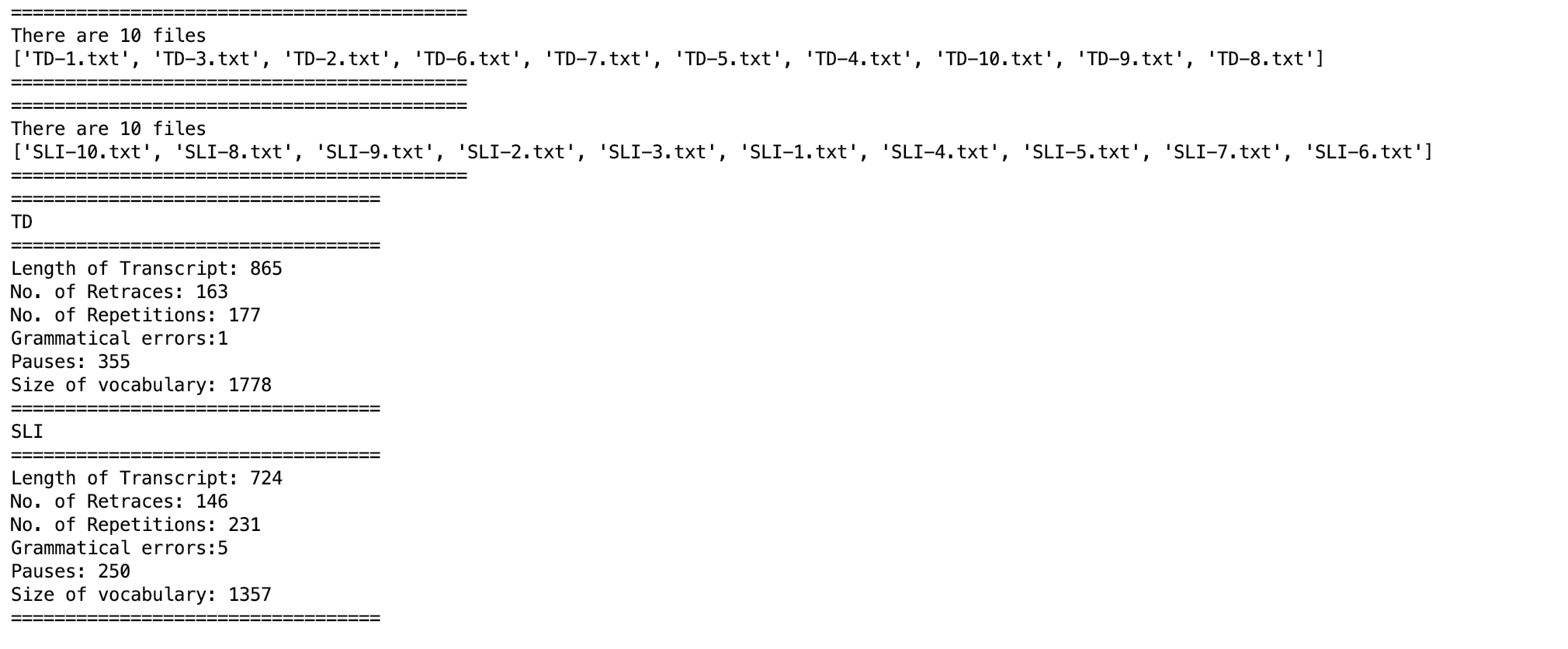
c. pip install matplotlib

1. Once all these are installed, we can proceed ahead to run the scripts. Run the scripts by using the command: ' **python task1.py**'. Run the remaining tasks similarly by using the similar command by replacing just the script name.
2. The output of the script if successful will look as below:

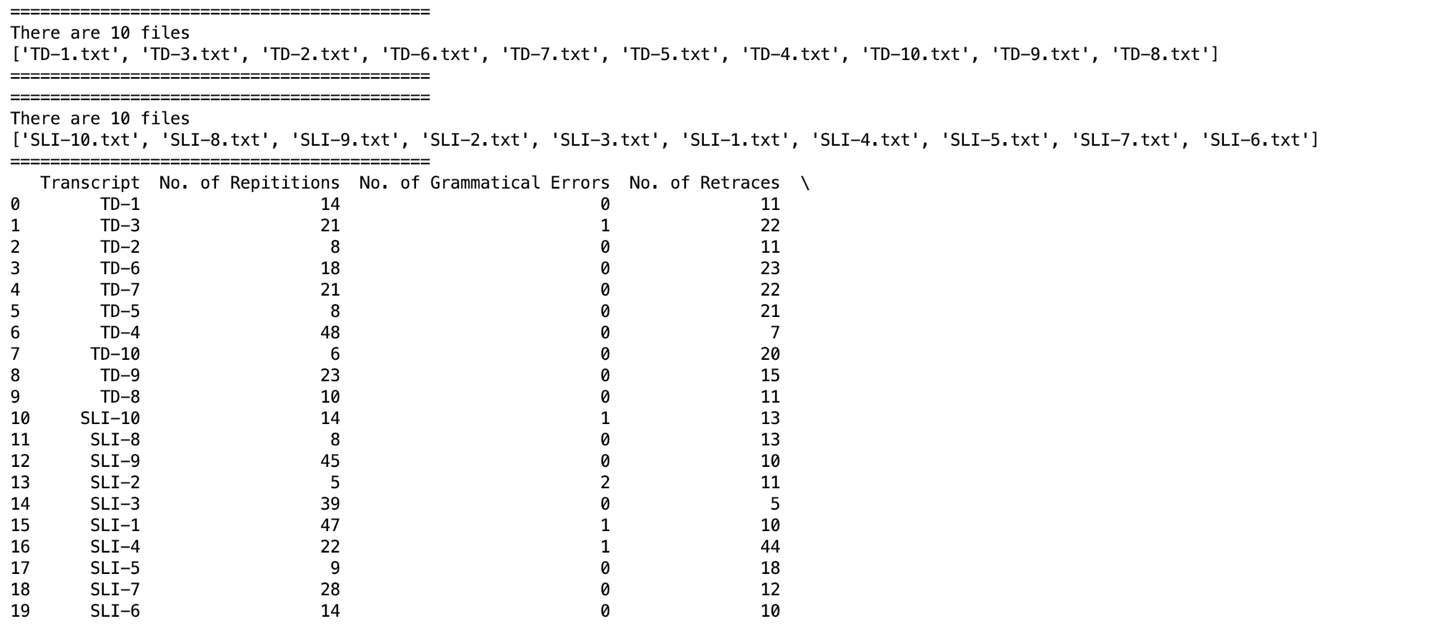
Screenshot 1 Output of Task-1

1. When the ENNI dataset is not in the directory, task 1 will be unsuccessful and no files will be written. We assume that the user will comply to instructions 3,4 and 5 as mentioned above. 

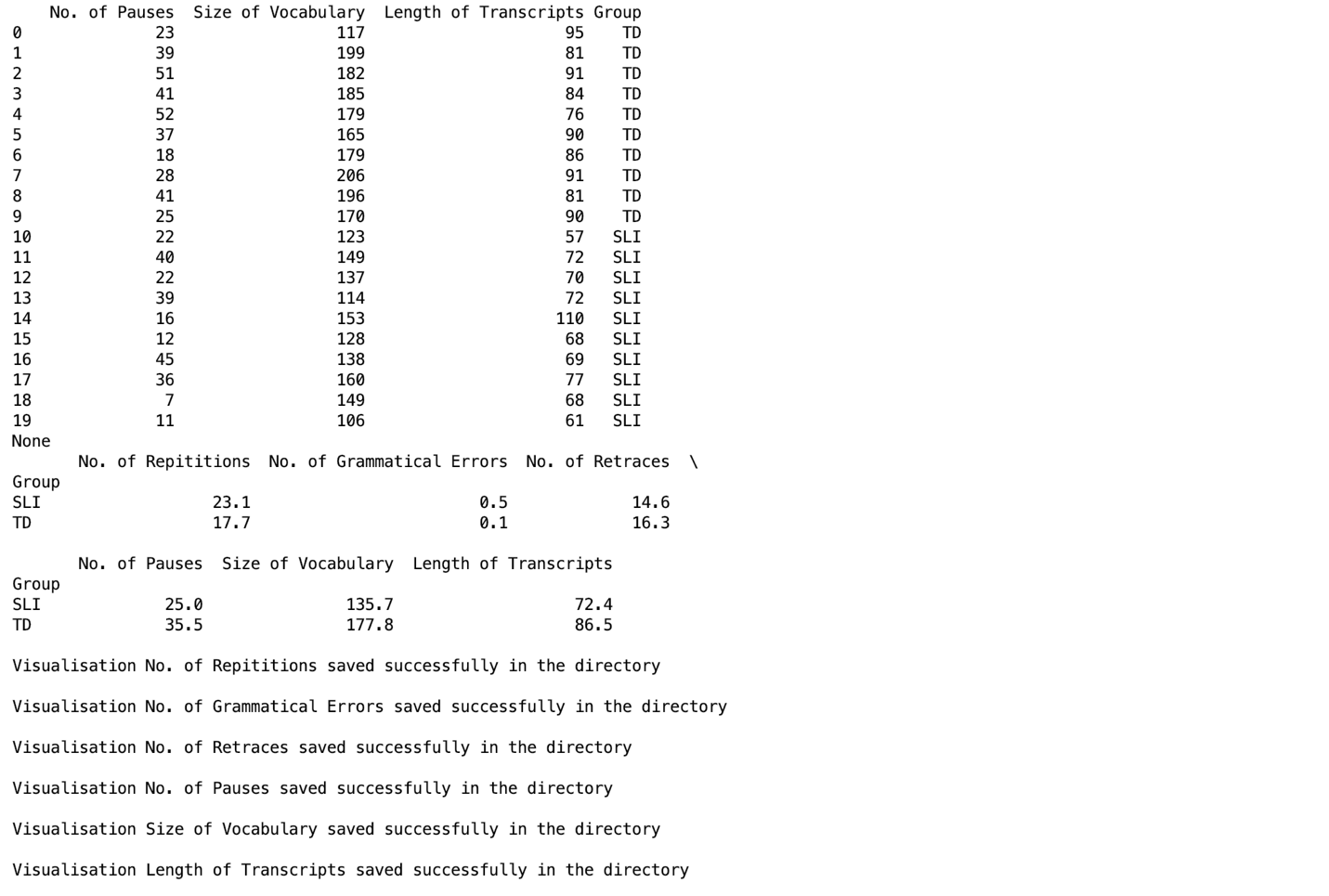
Screenshot 2 Unsuccessful output of Task-1

1. Task 2 requires Task 1 to be successful because we import variables and functions from Task-1.
2. Below is the output when the Task 2 is run:

Screenshot 3 Output of Task-2

1. Similarly, for task 3 we assume that task 2 was successfully executed and the output for the task will be:

Screenshot 4 Output of Task 3 (a)



Screenshot 5 Output of Task-3 (a)

Below is the sample image that is saved:

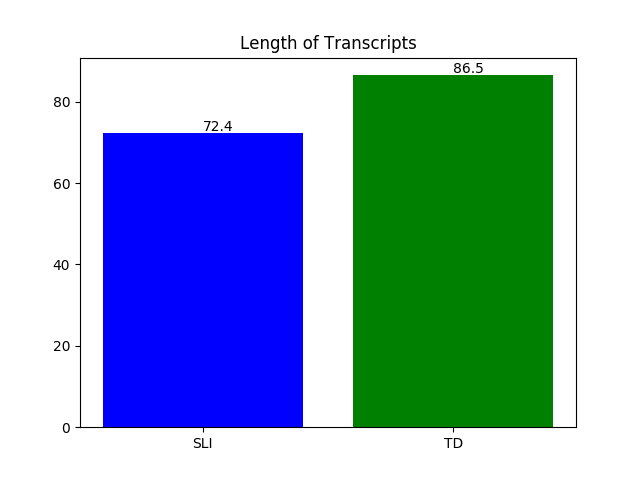


Figure 1 Saved image from Task-3

**Limitations**

1. The code cannot handle file in different folder from working directory. Also, the naming conventions of should comply with instructions 3,4 and 5.
2. There should be exactly 20 files in total inside TD folder and SLI folder combined.
3. Task 2 will fail if Task 1 was unsuccessful. Likewise, task 3 will fail if the previous two tasks were not successful
4. The visualisation statistics do not create a single graph to compare all the statistics
5. User defined output path for the statistics is not possible for the scripts.

**Code Support**

1. Reading and writing a file. (Kroeger, FIT9133- Tutorial week 10, 2018)
2. Use of commands *'os.listdir'*, *'os.makedirs'* and *'os.getcwd'* (Ciborowski & Conrad, 2008) (Pieters, 2014)
3. Creating dataframe from lists and dictionaries. (Mofitt, 2016)
4. Appending columns to a data frame based on condition. (unutbu & smci, 2013)
5. Plotting multiple bar graphs with categorical data. (tsabsch, 2018)
6. Plotting labels on the bar graph (Icastillov, 2016)
7. Widening the output in the python console. (Jacob & jezrael, 2015)

# References

alecxe. (2014, April 10). *Exception Handling in Python (Try…Except)*. Retrieved August 10, 2018, from stackoverflow: https://stackoverflow.com/questions/22990451/exception-handling-in-python-try-except

Ciborowski, D., & Conrad, B. (2008, November 07). *How can I safely create a nested directory in Python?* . Retrieved from stackoverflow: https://stackoverflow.com/questions/273192/how-can-i-safely-create-a-nested-directory-in-python

Icastillov. (2016, November 08). *How to write text above the bars on a bar plot (Python)?* Retrieved October 11, 2018, from stackoverflow: https://stackoverflow.com/questions/40489821/how-to-write-text-above-the-bars-on-a-bar-plot-python

Jacob, K. J., & jezrael. (2015, October 27). *Python pandas, how to widen output display to see more columns?* Retrieved October 11, 2018, from stackoverflow: https://stackoverflow.com/questions/11707586/python-pandas-how-to-widen-output-display-to-see-more-columns

Kroeger, G. (2018, August 8). Python Tutorial. *FIT9133*. Melbourne.

Mike\_R. (2017, June 2). *Tic Tac Toe: display score, prevent duplicate entries, print draw*. Retrieved August 8, 2018, from stackoverflow: https://stackoverflow.com/questions/44325584/tic-tac-toe-display-score-prevent-duplicate-entries-print-draw

Mofitt, C. (2016, September 06). *Practical Business python*. Retrieved October 03, 2018, from pbpython: http://pbpython.com/pandas-list-dict.html

Petri, U., & Gutmann, H. (2017, April). *PyFormat*. Retrieved August 8, 2018, from PyFormat: https://pyformat.info/

Pieters, M. (2014, March 5). *Python - how to find files and skip directories in os.listdir*. Retrieved September 28, 2018, from stackoverflow: https://stackoverflow.com/questions/22207936/python-how-to-find-files-and-skip-directories-in-os-listdir

Purcaru, G. (2010, November 12). *How to read two inputs separated by space in a single line?* Retrieved August 8, 2018, from stackoverflow: https://stackoverflow.com/questions/4162815/how-to-read-two-inputs-separated-by-space-in-a-single-line

tsabsch. (2018, January 05). *Plot multiple bars for categorical data*. Retrieved October 03, 2018, from stackoverflow: https://stackoverflow.com/questions/48157735/plot-multiple-bars-for-categorical-data

unutbu, & smci. (2013, November 11). *Pandas conditional creation of a series/dataframe column*. Retrieved October 05, 2018, from stackoverflow: https://stackoverflow.com/questions/19913659/pandas-conditional-creation-of-a-series-dataframe-column