# What can programming language do for our analysis?: The usefulness of Jupyter Notebook in the analysis of language testing

Yusuke Kondo[1] and Yutaka Ishii[2]

[1]Global Education Center, Waseda University,

[2]Center for Higher Education Studies, Waseda University

yusukekondo@waseda.jp, yutakaishii@aoni.waseda.jp

## Abstract

In this paper we describe the usefulness of Jupyter Notebook and R language in the analysis of learner language and statistical analysis.

## Keywords

Leaner language analysis, Jupyter Notebook, R

## Introduction

In the field of language testing, we may need to analyze learners' speech or writing in order to examine property of the test that we make. Corpus analysis software, software for finding word patterns, is very useful in this kind of analysis. However, sometimes, we may encounter some difficulty on that kind of software, because we may need to analyze an aspect of learners' performance that the software cannot analyze. In that case, we write codes to analyze such aspect of learners' performance in a programming language, such as Perl or Python. On the other hand, in the analysis of language testing, we always need to analyze test score or evaluation score. In so doing, we need some help of some statistical software, but we cannot know how the statistical software calculate because we just enter our data into the software and click on some buttons. Therefore, that makes us to use R, a programming language for statistical computing. As a consequence, we are obliged to use two kinds of environment, one is for text processing of learner language and the other, for statistical computing, which leads to inefficiency in our analysis. In this paper, in order to propose a solution to this problem, we demonstrate the usefulness of Jupyter Notebook, which a browser-based computational environment for Python. In Jupyter Notebook, a language-agnostic environment, we can add support for R, which means that both statistical computing and text processing are available in a single environment.

## 1 Example

The example used here is an analysis of speech and evaluation data in a story retelling speaking test (Koizumi and Hirai, 2010). The authors are trying to introduce automated scoring system to this test.

The data that we obtained were the transcriptions of 50 responses and their evaluations by teachers. In this test, firstly, the test-takers read the story below in two minutes, answer three comprehension questions in 40 minutes, and retell the story that they read in two minutes. The instructions on retelling are: Read the passage and retell the story as closely as possible. The test-takers are also given some keywords, some target grammatical items and some notes on vocabulary. Below is the text to analyzed here.

**Taking violin lessons**

When Rena was a little girl, she enjoyed taking violin lessons, but after a while, she was getting tired of it. Her teacher said that she should practice playing the violin every day. Rena's parents told her that if she didn't practice harder, she would have to stop taking lessons. They said that the lessons were too expensive to continue if she was going to be lazy. In the end, Rena quit her violin lessons.

Several years later, Rena went to a concert by a famous violinist with her parents. She was impressed by the beautiful sound of the music. So she decided to start playing the violin again. Now,

she never misses a day of practice. She hopes one day to be able to play like the famous violinist.

The retellings were evaluated based on three levels:

- Gives the outline of the story not including details
- Gives the outline of the story including some details
- Gives the outline of the story including almost every detail

What we want to do in the data is to find characteristics of leaner language that can predict the evaluation scores.

## 2 Analysis of learner language

The retellings were transcribed by the authors and a graduate student who are majoring in Applied Linguistics. Below is an example of the transcriptions.

when rena was nnn rena was a girl she started taking lesson taking violin lesson but she not for a while after a while she was tired it her teacher said that her she practice everyday so uhh uhh her parents said that so her parents told that the lesson was too expensive continually if so she quit leaning violin several years later uhh she go she nnn she went to the concert with her parents when she heard beautiful music of violin she impressed uhh by this music uhh she she start again learning violin her hope is one day she play violin beautiful likes she heard

In the transcriptions, capital letters, commas, and periods were not used, and fillers, false starts, repetitions are transcribed.

Because we try to find the similarities between the contents of the test-takers retellings and that of the target passage, we may delete the dysfluencies, such as fillers and repetitions. In order to do this, we can use Python, a script language on Jupyter Notebook. By writing several lines in Python, we can delete the dysfluencies, and can count the tokens and the types of the text. Furthermore, we can easily measure cosine similarity distance, Jaccard index, and Euclidian distance by adopting vector space model.

## 3 Statistical analysis of evaluation scores

A variety of packages are available in Python, but there is no package in python to calculate indices of psychometrics, such as inter-rater reliability. Therefore, we need other software or environment. However, in Jupyter Notebook, we can run R. Therefore, we can analyze learner language and evaluation scores in a single application.

**References**

Koizumi, R., & Hirai, A. (2010). Exploring the quality of the story retelling speaking test: Roles of story length, comprehension questions, keywords, and opinions. *Annual review of English language education in Japan 21*, 211-220.

Jupyter Project (n.d). Jupyter Notebook version 4.1.1. Available at http://jupyter.org

Python Software Foundation (n.d.). Python Language Reference, version 2.7. Available at http://www.python.org .

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/.