

# Exploring Roadblocks to Second Language Acquisition

Brian Diep

09/12/2020

## Keywords

R Markdown, Language Education, Learning

## Abstract

Second language education has typically been confined to a traditional classroom experience with a teacher. However with the advent of the Internet and accessible technology, digital language education has been immensely successful in part due to its scale and flexibility. However, one of the major benefits of digitizing language education is that it is now possible to collect data in real-time and simultaneously use it to improve the learning experience for users. It is therefore important that we analyze the relationship between user-model interactions, aspects of language, and language acquisition to improve language acquisition for all. We approach this problem by using data gathered from the popular language education app Duolingo, we seek to model any trends between using Bayesian linear regression. Through this methodology we find that (blah blah blah)

## Introduction

Learning a second language is a difficult task for most adults who have passed the critical stage of language acquisition as children. However, that has not stopped second language education from being a core part of education systems across the world. Unfortunately, these systems often leave much to be desired as they lack the flexibility and scope necessary to truly learn a language. Outside of the traditional system of language education, the widespread proliferation of technology such as the Internet, personal computing, and mobile devices have allowed other actors to break into the language education, in particular mobile apps. One of the greatest benefits that these apps have is that they can take user performance metrics and use them to guide decisions on teaching users.

In this report, we seek to explore this data driven approach using the free language education app Duolingo as a case study. The app has roughly 120 million users and is an example of one of the most popular of such language education apps available. We specifically focus on Duolingo's usage of spaced repetition as a method for teaching its users various grammar points. We also seek more broadly to identify any consistent roadblocks for second language learners that can be improved.

The remaining report will continue to further explore the data sets and methodologies used to gather the data, an analysis of the model used, results gathered from. In particular. Our findings suggest that (fasd-jflkasdjflksadf). Furthermore, we discuss and investigate relationships between our predictor and response variables followed by a discussion of the strengths and weaknesses of the model, and further avenues of study.

```
unique(data$ui_language)
```

```
## [1] "en" "pt" "es" "it"
```

```
unique(data$learning_language)
```

```
## [1] "de" "es" "pt" "en" "fr" "it"
```

Exploratory Data Analysis

Model

Results

Discussion

Limitations and Weaknesses

Future Work

Appendix

References