Advice on Statistical Methods for Mourad Abdennebi’s Research

STAT 688 Statistical Consulting with Dean Billheimer

Sasha Berliner (aiberliner@arizona.edu)

Shrey Arora (shreyarora@arizona.edu)

Kamal Singh (kamaljeetsingh@arizona.edu)

September 11, 2023

# Executive Summary

Mourad Abdennebi is a PhD student in second language acquisition and teaching. His research question investigates if participants can learn L2 nouns and prepositions with gestures better than non-gestures. Abdennebi provided de-identified data in an excel sheet and wanted advice on which statistical method to use for his analysis as well as which software can be used to perform the analysis. We recommend and demonstrate using R software to analyze simulated data and result interpretation. After constructing a generalized linear mixed model, analysis of the simulated data shows no significant difference between learning L2 words using gestures and no gestures, as well as between learning L2 nouns and prepositions.

# Detailed Summary

## Background

An L2 word is a word in the language one is learning, while an L1 word is a word in the participant’s native language. Abdennebi designed a survey where participants had to learn 18 nouns and 18 prepositions in L2 by providing them with an L1 translation. For 18 of the 36 words in total, consisting of 9 nouns and 9 prepositions, the participants were given a video of Abdennebi saying the L2 word, followed by a gesture of the word. For the other 18 words, the participants had the same video without the gesture. The end of the surveys tested the participant’s recollection by asking them to provide English translations to the 36 L2 words. The 18 gesture and 18 non-gesture words were the same in each survey, including the order in which they were asked. For every participant, Abdennebi recorded a 1 for a word the participant correctly recalled, and a 0 if incorrect or blank.

## Methods

We simulated a data set, with surveys A and B (the score pattern is similar to the original data). Resources to install and get started with R can be found in “An Introduction to R” (A.1).

Before we start with the data analysis, please refer to our email on 09/20/2023 to make necessary edits to your data file. Steps to analyze the data:

1. We recommend that merging of the files be done inside the statistical software so that we preserve the original data file in case we have to go back to it.
2. Combine survey versions A and B into one dataframe using the code given in A.2.
3. Before we begin with the actual analysis, we will reorganize the data. We created a new variable GA\_PS in the new\_data dataframe to get classifications such as “gestured\_nouns,” “nongestured\_nouns,” “gestured\_prepositions,” or “nongestured\_prepositions”, as you can refer to in code block (1) in the appendix.
4. Next, we visualize the data. Referring to code block (2), we summarize the data with grouped bar charts (bar plots) to visualize the proportions of correct responses for different combinations of “participant” and “GA\_PS” (a variable created in the previous code section). The group\_by(participant, GA\_PS) first groups the new\_data dataframe by both “participant” and “GA\_PS.” Within each combination of “participant” and “GA\_PS,” the code calculates the proportion of correct responses by summing number of rows where “score” equals 1, indicating a correct response, and divides the sum by the total number of rows in that combination. Find the resulting plots in the Results section. Referring to code block (3), we then visualize the data by plotting the distribution of scores for each treatment factor (gesture, no gesture, noun, preposition) against the proportion of correct scores.
5. Lastly, we recommend using a generalized linear mixed model so that we can test if there are any significant differences in learning L2 words with or without gestures and if they are nouns or prepositions. We chose this model because our response variable is binary (0 or 1, that is, incorrect or correct answer) in addition to both fixed and random effects of our treatment variables. Referring to code block (4), we fit a Generalized Linear Mixed Model (GLMM) and then summarize the results of the model.
6. In code block (5), we transform the log odds coefficient estimates produced in block (4) by exponentiation to get more interpretable odds ratio estimates. This does not change the significance of the results, just interpretability.

## Results

We acknowledge that the next few figures are probably not publication worthy. Rather, we suggest them as a visual means of familiarizing the investigator with the data and for evaluating whether the measurement tool is working as intended.

Figure 1, as constructed in code block (1), shows a plot of the proportion of correct responses for each participant given each possible combination of treatments: nouns with gestures, prepositions with gestures, nouns without gestures, and prepositions without gestures. It appears that most participants correctly got the nouns that were not accompanied by gestures, while most had difficulty with the nouns accompanied with gestures. It appears that gestures do not play a role in correctly recalling the prepositions.

A graph with colorful lines

Description automatically generated

Figure 1

It may be easier to assess the results if we create two separate plots, one for gestures versus no gestures, and one for nouns versus prepositions. See figures 2 and 3, corresponding to code block (2). From the plots alone, it appears that there is no difference between learning words with gestures and without gestures, but that it was easier for participants to remember nouns over prepositions.

A graph with red and blue lines

Description automatically generatedA graph with red and blue lines

Description automatically generated

Figures 2,3

In Figure 4, corresponding to code block (3), we visualize the data by plotting the distribution of scores for each treatment factor (gesture, no gesture, noun, preposition) against the proportion of correct scores. These plots merely show us the general trend of scores for each treatment factor. From these plots it appears that there is a higher proportion of correct answers when gestures were used, while the trend for proportion of correct answers for nouns and prepositions appear around the same.

A graph with colorful bars and text

Description automatically generatedA graph with red and blue bars

Description automatically generated

A graph of different colored bars

Description automatically generatedA graph of different shapes

Description automatically generated with medium confidence

Figure 4

## Results

The summary table of GLM model is given below:

A screen shot of a computer code

Description automatically generated

Participants and L2 words themselves may be contributing to random variation in this study and hence are considered as random effects in this model. And fixed effects here are the variables of interest accompanied by gestures and parts of speech. Looking at the Fixed effects section of the summary output from the GLMM model above, we have a column, labeled Pr(>|z|), of p-values corresponding to our treatment factors: gesture versus non-gesture, nouns versus prepositions, and the interaction between the two. In this project, the null (default) hypothesis is that gestures or parts of speech are not significantly related to learning an L2 word. The p-value tells us the probability of obtaining a test result that confirms the null hypothesis. Thus, a small p-value means that it would be unlikely the null hypothesis is true. However, for this simulated dataset we have p-values much greater than standard values of 0.05, so we fail to reject the null hypothesis and conclude that gestures or parts of speech are not significantly related to learning an L2 word.

Moreover, we observe two types of effects: random effects and fixed effects. The coefficients for fixed effects are not presented in terms of probabilities but rather on a different scale known as the log-odds or logit scale. The logit transformation converts values originally between 0 and 1 (representing probabilities) into a different range, spanning from -inf to +inf.

The coefficients for log odds are given below along with their confidence intervals:

A close-up of a computer screen

Description automatically generated

After that we find the Confidence intervals of all the variables and their interactions. The code for finding the confidence intervals is provided in the appendix (code block 5a).

Since all these coefficients and their intervals are on log odds, we need to exponentiate the estimates and confidence interval levels to get the results for odds ratios (code block 5b).

The estimated coefficients and their confidence intervals after exponentiating them are given below:

A close-up of a white background

Description automatically generated

**Interpretation of Exponentiated Coefficients:**

When we increase any variable (say accompanied\_by\_gesture) by one unit (from no to yes) the odds are multiplied by the exponentiated coefficient of accompanied\_by\_gesture, while holding other variables constant. If the exponentiated coefficient is equal to 1, it suggests that there is no effect of the predictor on the odds of the scoring 1 (learning). If the exponentiated coefficient is greater than 1, it suggests that an increase in the predictor variable is associated with higher odds of scoring 1 (learning). If the exponentiated coefficient is less than 1, it suggests that an increase in the predictor variable is associated with lower odds of the scoring 1 (learning).

**Interpretation of Exponentiated Confidence Intervals:**

Confidence Interval excludes 1: If the CI is either above 1 or below 1 (but not including 1), it suggests that the odds ratio (OR) or exponentiated coefficient is statistically significant. This means that there is evidence that the predictor variable has a significant effect on the odds of the event occurring.

Confidence Interval Includes 1: If the CI includes 1, it suggests that the OR or exponentiated coefficient is not statistically significant. In this case, there is insufficient evidence to conclude that the predictor variable has a significant effect on the odds of the event occurring.

**Result interpretation**

Since confidence intervals for predictor variables ‘accompanied by gesture’, ‘part of speech’, and their interaction term all include 1 in their respective confidence intervals, we conclude here that in this simulated dataset there is no significant difference in leaning the L2 words with vs without gestures as well as no significant difference in leaning the L2 words as nouns vs prepositions. Importantly, it is acceptable and sometimes preferable to describe the direction of the association found in the sample data, even when the results are not statistically significant, i.e., when we cannot infer a difference beyond the sample. For example, we can say that in this experiment, there was slightly increased odds of words being identified when accompanied by gestures as compared to when the words were not accompanied by gestures, but this increase was not statistically significant (OR: 1.03; 95% CI: 0.70 - 1.51), and a slightly decreased odds of prepositions being identified as compared to nouns, which was similarly not statistically significant (OR: 0.99; 95% CI: 0.57 - 1.46).

If, however, say the confidence interval for ‘accompanied by gesture’ was above 1 (and excluded 1) e.g., 95% CI: 1.0159 – 1.5099 with odds ratio 1.03, we could say that the odds of learning an L2 word with gestures is 1.03 times higher than the odds of learning an L2 word without gestures.

## Appendix

A.1 Book Introduction to R can be found at <https://intro2r.com/index.html>

A.2 Read in excel file and combine surveys.

# install packages & load in your libraries required for data processing

install.packages("readxl") # required to import excel sheets in R

install.packages("janitor") # used to clean and standardize variable names

install.packages("tidyverse") # used to manipulate data

install.packages("lme4") # required to run the GLMM

library(readxl)

library(janitor)

library(tidyverse)

library(lme4)

# importing data & merging surveys

sheet\_a <- read\_xlsx("YOUR\_FILE\_PATH/Gestures Study\_Scoring responses.xlsx",

sheet = "A") %>%

clean\_names()

sheet\_b <- read\_xlsx("YOUR\_FILE\_PATH /Gestures Study\_Scoring responses.xlsx",

sheet = "B") %>%

clean\_names()

# all variable names (column names) in both sheets should be the same to combine

combined\_data <- rbind(sheet\_a,sheet\_b)

(1)

1. group\_by(participant): This part of the code groups the combined\_data dataframe by the “participant” variable, meaning that subsequent operations will be applied within each participant group separately.
2. mutate(GA\_PS = ifelse(...)): Within each participant group, this code uses the mutate function to create a new variable GA\_PS. The value of GA\_PS is determined by nested ifelse statements based on conditions involving the variables “accompanied\_by\_gesture” and “part\_of\_speech.” It assigns values like “gestured\_nouns,” “nongestured\_nouns,” “gestured\_prepositions,” or “nongestured\_prepositions” to GA\_PS based on the conditions. If none of the conditions are met, it assigns NA to GA\_PS.
3. ungroup(): After creating the new variable, this code uses the ungroup function to remove the grouping by “participant” so that subsequent operations will be applied to the entire dataset, not just within participant groups.
4. select(participant, accompanied\_by\_gesture, part\_of\_speech, GA\_PS, score): Finally, this code selects specific columns from the new\_data dataframe, including “participant,” “accompanied\_by\_gesture,” “part\_of\_speech,” the newly created “GA\_PS” variable, and “score.” It creates a new dataframe called new\_data containing only these selected columns.

In summary, this code block (1) groups the original combined\_data by “participant,” creates a new variable GA\_PS based on conditional logic, ungroups the data, and then selects specific columns to create a new dataframe new\_data with the desired columns and the newly created variable.

# creating combined categories  
  
new\_data <- combined\_data%>%   
 group\_by(participant) %>%   
 mutate(GA\_PS = ifelse(accompanied\_by\_gesture == "yes" & part\_of\_speech == "noun", "gestured\_nouns",   
 ifelse(accompanied\_by\_gesture == "no" & part\_of\_speech == "noun", "nongestured\_nouns",  
 ifelse(accompanied\_by\_gesture == "yes" & part\_of\_speech == "preposition", "gestured\_prepositions",  
 ifelse(accompanied\_by\_gesture == "no" & part\_of\_speech == "preposition", "nongestured\_prepositions",NA))))) %>%   
 ungroup() %>%   
 select(participant, accompanied\_by\_gesture, part\_of\_speech, GA\_PS, score)

(2)

Measuring the proportion of correct responses, we plot a bar chart grouped into 4 categories: Nouns with gestures, prepositions with gestures, nouns without gestures, prepositions without gestures.

new\_data %>%   
 group\_by(participant,GA\_PS) %>%   
 summarise(Proportion\_Correct = sum(score == 1)/length(score)) %>%   
 ggplot(aes(x= participant, y = Proportion\_Correct, fill = GA\_PS)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 theme\_minimal()

## `summarise()` has grouped output by 'participant'. You can override using the  
## `.groups` argument.

Plot a bar chart grouped into two categories: gestures and no gestures.

new\_data %>%   
 group\_by(participant,accompanied\_by\_gesture) %>%   
 summarise(Proportion\_Correct = sum(score == 1)/length(score)) %>%   
 ggplot(aes(x= participant, y = Proportion\_Correct, fill = accompanied\_by\_gesture)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 theme\_minimal()

## `summarise()` has grouped output by 'participant'. You can override using the  
## `.groups` argument.

Plot a bar chart grouped into two categories: nouns and prepositions.

new\_data %>%   
 group\_by(participant,part\_of\_speech) %>%   
 summarise(Proportion\_Correct = sum(score == 1)/length(score)) %>%   
 ggplot(aes(x= participant, y = Proportion\_Correct, fill = part\_of\_speech)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 theme\_minimal()

## `summarise()` has grouped output by 'participant'. You can override using the  
## `.groups` argument.

(3)

Measuring the proportion of correct answers, we plot by distribution of scores.

This bar plot is categorized by nouns with gestures, prepositions with gestures, nouns without gestures, and prepositions without gestures.

new\_data %>%   
 group\_by(participant, GA\_PS) %>%   
 summarise(Proportion\_Correct = sum(score == 1)/length(score)) %>%   
 ggplot(aes(x = Proportion\_Correct, fill = GA\_PS)) +  
 geom\_histogram(binwidth = 0.05, position = "dodge") +  
 labs(x = "Category",  
 y = "Proportion Correct") +  
 theme\_minimal()

## `summarise()` has grouped output by 'participant'. You can override using the  
## `.groups` argument.

Bar plot categorized by gestures versus no gestures:

new\_data %>%   
 group\_by(participant, accompanied\_by\_gesture) %>%   
 summarise(Proportion\_Correct = sum(score == 1)/length(score)) %>%   
 ggplot(aes(x = Proportion\_Correct, fill = accompanied\_by\_gesture)) +  
 geom\_histogram(binwidth = 0.05, position = "dodge") +  
 labs(x = "Category",  
 y = "Proportion Correct") +  
 theme\_minimal()

## `summarise()` has grouped output by 'participant'. You can override using the  
## `.groups` argument.

Bar plot categorized by noun versus preposition:

new\_data %>%   
 group\_by(participant, part\_of\_speech) %>%   
 summarise(Proportion\_Correct = sum(score == 1)/length(score)) %>%   
 ggplot(aes(x = Proportion\_Correct, fill = part\_of\_speech)) +  
 geom\_histogram(binwidth = 0.05, position = "dodge") +  
 labs(x = "Category",  
 y = "Proportion Correct") +  
 theme\_minimal()

## `summarise()` has grouped output by 'participant'. You can override using the  
## `.groups` argument.

Four bar charts in one plot, each by its own category:

new\_data %>%   
 group\_by(participant, GA\_PS) %>%   
 summarise(Proportion\_Correct = sum(score == 1)/length(score)) %>%   
 ggplot(aes(x = Proportion\_Correct)) +  
 geom\_histogram(binwidth = 0.05, position = "dodge") +  
 facet\_wrap(~GA\_PS) +  
 labs(x = "Proportion Correct",  
 y = "Frequency") +  
 theme\_minimal()

## `summarise()` has grouped output by 'participant'. You can override using the  
## `.groups` argument.

(4)

GLMM:

model<-glmer(score ~ accompanied\_by\_gesture + part\_of\_speech + accompanied\_by\_gesture\*part\_of\_speech + (1|participant) + (1|item\_l2), family = binomial, data = combined\_data)

summary(model)

Code explanation of (4):

1. The model formula specifies that the binary outcome variable score is being regressed on the predictors accompanied\_by\_gesture, part\_of\_speech, and their interaction (accompanied\_by\_gesture \* part\_of\_speech). This formula is used to model the relationship between the predictors and the binary outcome variable, assuming a binomial distribution (specified by family = binomial).
2. The items participant and item\_l2 are random effects terms. Since we selected a random sample of participants from a population of L1 speakers, and since we also selected a random sample of words from a population of possible L2 words, we are using participants and items as the random effects terms.

(5a)

# referred https://stats.oarc.ucla.edu/r/dae/mixed-effects-logistic-regression/ to get the confidence intervals   
  
se <- sqrt(diag(vcov(model)))  
# table of estimates with 95% CI  
(tab <- cbind(Est = fixef(model), LL = fixef(model) - 1.96 \* se, UL = fixef(model) + 1.96 \*  
 se))

(5b)

exp(tab)