Analyze the Data

Erin M. Buchanan

Last Update 2021-08-25

Libraries

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
library(psych)
library(reshape)
```

```
##
## Attaching package: 'reshape'

## The following object is masked from 'package:dplyr':
##
##     rename
library(tidyr)
```

```
##
## Attaching package: 'tidyr'

## The following objects are masked from 'package:reshape':
##
##     expand, smiths
library(ggplot2)
```

```
##
## Attaching package: 'ggplot2'

## The following objects are masked from 'package:psych':
##
##     %+%, alpha
```

Import the Files

```
targets_only <- read.csv("../05_Data/output_data/prime_data.csv")
targets_only_no2.5 <- read.csv("../05_Data/output_data/prime_data_no2.5.csv")
targets_only_no3.0 <- read.csv("../05_Data/output_data/prime_data_no3.0.csv")
```

Merge All Languages

This demonstration shows processing of an English related demonstration. Each language will be stored separately and processed separately. We will then combine the languages together for Hypothesis testing. For this demonstration of the code, we will simply show English. This process will be repeated for data without 2.5 z-score trials and 3.0 z-score trials.

```
# merge all targets_only files
targets_only$lang <- "English"
targets_all <- targets_only # for this demo
```

Hypothesis 1

Is semantic priming a non-zero effect? To assess this research question, we will examine the confidence interval of the semantic priming effect to determine if the lower limit of the confidence interval is greater than zero. Therefore, Hypothesis 1 is for semantic facilitation with faster response latencies for related words in relation to unrelated words. We will provide estimates for all languages and each language individually.

```
# calculate all languages together
all_language <- psych::describe(targets_all$avgZ_prime)
all_language
```

```
##    vars n mean   sd median trimmed  mad   min  max range skew kurtosis   se
## X1    1 3  0.4 1.61  -0.18     0.4 0.96 -0.83 2.23  3.06 0.32    -2.33 0.93
# lower
all_language$mean - qt(.025, nrow(targets_all), lower.tail = F)*all_language$se
```

```
## [1] -2.55505
# upper
all_language$mean + qt(.025, nrow(targets_all), lower.tail = F)*all_language$se
```

```
## [1] 3.36245
# by language
each_language <- psych::describeBy(targets_all$avgZ_prime, group = targets_all$lang)
each_language
```

```
##
##  Descriptive statistics by group
## group: English
##    vars n mean   sd median trimmed  mad   min  max range skew kurtosis   se
## X1    1 3  0.4 1.61  -0.18     0.4 0.96 -0.83 2.23  3.06 0.32    -2.33 0.93
# lower
each_language$mean - qt(.025, nrow(targets_all), lower.tail = F)*each_language$se
```

```
## numeric(0)
# upper
each_language$mean + qt(.025, nrow(targets_all), lower.tail = F)*each_language$se
```

```
## numeric(0)
```

Hypothesis 2

How does the semantic priming effect vary across languages? We will compare the confidence intervals of the semantic priming effect for each language to determine if they overlap. The effect will be considered equal when confidence intervals overlap between languages and different when they do not overlap. In Hypothesis 2, we do not pick a direction for the effects, but expect differences between languages.

For this hypothesis, we will use the analyses provided above, but interpret the confidence intervals. We will create visualizations for this hypothesis.

```
# example plot
ggplot(targets_all, aes(lang, avgZ_prime)) +
  stat_summary(fun.data = mean_cl_normal,
               geom = "errorbar",
               width = .2,
               position = position_dodge(width = .2)) +
  stat_summary(fun = mean,
               geom = "bar",
               position = position_dodge()) +
  xlab("Language") +
  ylab("Z Score Priming") +
  theme_classic()
```