# Analysis of the effects of context awareness and experience taking the SPEAK test on listener perceptions of accented speech

**Client:** Amber Franklin, Ph.D. CCC-SLP

**Group Members:** Tyler Sagendorf, Yu Zhang, Nannan Zhao

# 1. Executive Summary

## 1.1 Background

The purpose of this study is to determine which factors, if any, influence listener perceptions of accented speech. This is important because these factors can significantly impact the lives of non-native speakers—namely, how they are treated by others due to perceived ability. Specifically, we want to know if experience taking the Speaking Proficiency English Assessment Kit (SPEAK) oral exam, awareness of the context of discourse, speaker proficiency, or any combinations of two of these factors influence listener perceptions of accented speech. The specific research questions are provided in section 2.3.

## 1.2 Methods

We created some preliminary plots in RStudio to examine the effects of SPEAK experience (SE), context awareness (CA), and proficiency (Prof) on the averages of the Accentedness, Comfort, and Comprehensibility scores. Additionally, we looked at plots of the interactions between SE and CA, SE and Prof, and CA and Prof. Finally, we fit an analysis of variance (ANOVA) model to determine which factors, if any, were important for predicting average scores.

## 1.3 Results

Based on our analysis, there appears to be evidence that a speaker's proficiency (strong evidence) and the interaction between proficiency and SPEAK experience of listeners (some evidence) influence the average of the Accentedness, Comfort, and Comprehensibility scores. That is, these are the only factors that seem to influence listener perceptions of accented speech.

# 2. Introduction

## 2.1 Background

The overall purpose of this report is to determine how context awareness, listener experience, taking the SPEAK test, and speaking proficiency influence listener perceptions of accented speech (Cooley, 2019).

Previous studies have found that "listeners are more likely to assign lower ratings of status and credibility to non-native speakers," (Lev-Ari and Keysar, 2010; Lindemann, 2003 as cited in Cooley, 2019) and that "foreign-accented speakers are rated less competent, less hirable, and less suitable for jobs with high communication demands" (Hansen, Rakic, and Steffens, 2014; Hosoda and Stone-Romero, 2009 as cited in Cooley, 2019). Additionally, international teaching assistants (ITAs) may be assigned lower ratings in student evaluations due to the "presence of a foreign accent." The presence of an accent may similarly lead to "unfavorable promotion and tenure decisions" (Kang, Rubin, and Lindemann, 2014; Shao, Anderson, & Newson, 2007 as cited in Cooley, 2019). This means that the factors that influence a listener's perceptions of accented speech are important because they can have profound effects on the lives of non-native speakers.

This study was motivated by previous research that showed "general familiarity with non-native speech and familiarity with the topic of discourse (semantic context) facilitate comprehension of accented speech" (Gass and Varonis, 1984; Kennedy, & Trofimovich, 2008 as cited in Cooley, 2019). That is, context awareness improves comprehension and, therefore, listener perceptions of accented speech. There is also the idea that "manufactured contact and perspective-taking interventions are effective methods of influencing listener attitudes towards accented speech" (Kang, Rubin, & Lindemann, 2014; Manohar & Appiah, 2016 as cited in Cooley, 2019). In terms of this study, it would mean that SPEAK experience (understanding the perspective of the speakers by taking a different version of the same test) would also improve listener perceptions of accented speech.

For this study, 56 participants (mainly female undergraduate SPA majors) were selected to listen to 12 different audio files recorded by a Native Chinese speaker with low English speaking proficiency and a Native Chinese speaker with high English speaking proficiency. Each of the audio files were rated on three measures (accentedness, comprehensibility, and comfort level with having this speaker as an instructor) using a visual analog scale that could take on integer values from 100 to 900. In addition, the listeners were randomly assigned to one of four different treatment combinations (two factors, each with two levels)—context-aware (the listeners "received information related to the content of each stimulus presentation"), SPEAK experience (the listeners "completed a different form of the SPEAK test in their native language prior to rating the speech samples"), naive (the listeners "did not receive any contextual information" and they "did not complete the SPEAK test"), and both (context-aware and SPEAK experience). Each of the listeners heard both the high and low proficiency speakers in a random order as well (Cooley, 2019).

The experiment conducted for the study is a split-plot design with two whole plot factors (context awareness and SPEAK experience) and one subplot factor (speaker proficiency). The split-plot design is detailed in section 3.

## 2.2 Data

In RStudio, we took the average of the accentedness, comprehensibility, and comfort scores because this seemed to be an appropriate measure of "average perceptions of accented speech." Then, we grouped the original data by listener, context awareness, SPEAK experience, and proficiency. This allowed us to condense the data into 112 rows (2 average scores per listener—one for the high proficiency speaker and the other for the low proficiency speaker). The first 6 rows of this dataset are shown below, and descriptions of this and the original dataset are available in section 8.2.

| listener | CA | SE | prof | code | Avg_Score |
|---|---|---|---|---|---|
| B01 | YES | YES | High | B | 167 |
| B01 | YES | YES | Low | B | 538 |
| B02 | YES | YES | High | B | 350 |
| B02 | YES | YES | Low | B | 734 |
| B04 | YES | YES | High | B | 222 |
| B04 | YES | YES | Low | B | 726 |

## 2.3 Research Questions

These are all of the research questions that our groups answered. The research questions that we address in this report are in bold.

1. **How does experience taking a test of spoken English proficiency affect listener perceptions of accented speech?**
2. **How does context awareness affect listener perceptions of accented speech?**
3. **Is there an interaction between listener condition (Naïve, CA, SE, Both) and speaker proficiency?**
4. What is the intraclass correlation coefficient?
5. Do Accentedness and Comprehensibility predict Comfort?
6. **Is there a significant difference in the effects of context awareness and test-taking experience on listener perceptions of accented speech?**
7. **Does the combination of experience taking a test of spoken English proficiency and context awareness result in different ratings than either test-taking experience or context awareness alone? Essentially, is the "Both" group different from either the "SE" or "CA" groups?**

# 3. Methods

To answer the above research questions, we manually fixed data-entry errors[1] in the "SPEAK Results with Ciara's data 2019_9_10 with comments.xlsx" file and read it into RStudio. Then, we created a new variable by averaging the accentedness, comprehensibility, and comfort scores. From there, we condensed the original 1344-row data set into 112 rows (the overall average score for each of the two speakers from the 56 listeners). Next, we created some preliminary plots to visualize the effects of the variables (context aware, SPEAK experience, and proficiency) and their interaction effects. Finally, we fit a mixed model with these variables and

[1] Rows 323-337 had the values for Accentedness, Comprehensibility, and Comfort shifted over so that the first column had 2-digit numbers, the second had a 1-digit number followed by a space and a two-digit number, and the last had a 1-digit number followed by a space and a 3-digit number.

their 2-way interactions as predictors of the average scores in SAS. We also included the listener variable nested within the interaction between SPEAK experience and context awareness as a random factor; since the 56 listeners represent a sample from a larger population of listeners, the levels in this study (each listener) are a subset of all possible levels (Montgomery, 2017, p. 111).

This model was chosen because the experiment follows a split-plot design with two whole plot factors (context awareness and SPEAK experience) and one subplot factor (proficiency). A basic split-plot design consists of a whole plot factor (also known as the "hard-to-change" factor) and a subplot factor ("easy-to-change" factor). The reason we selected our whole plot factors to be context awareness (CA) and SPEAK experience (SE) is because the combination is held constant (fixed) for the change in proficiency (Prof) (Montgomery, 2017, p. 634-640).

Also, characteristic of split-plot designs are the two independent levels of randomization. First, the listeners were randomly assigned to each group. Second, the order in which the listeners heard the audio files was randomized. This randomization scheme tells us about the experimental units (the units—usually subjects—that receive the treatment) for the whole plot and the subplot—the whole plot experimental units are the different listeners and the subplot experimental units are the sets of 12 audio files from the two speakers (Cooley, 2019; Montgomery, 2017).

The linear model (Montgomery, 2017, p. 635-636) for this split-plot design is

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \nu_{l(ij)} + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \varepsilon_{l(ijk)},$$

$$\text{with } i = 1,2;\ j = 1,2;\ k = 1,2;\ l = 1,2,\ldots 14$$

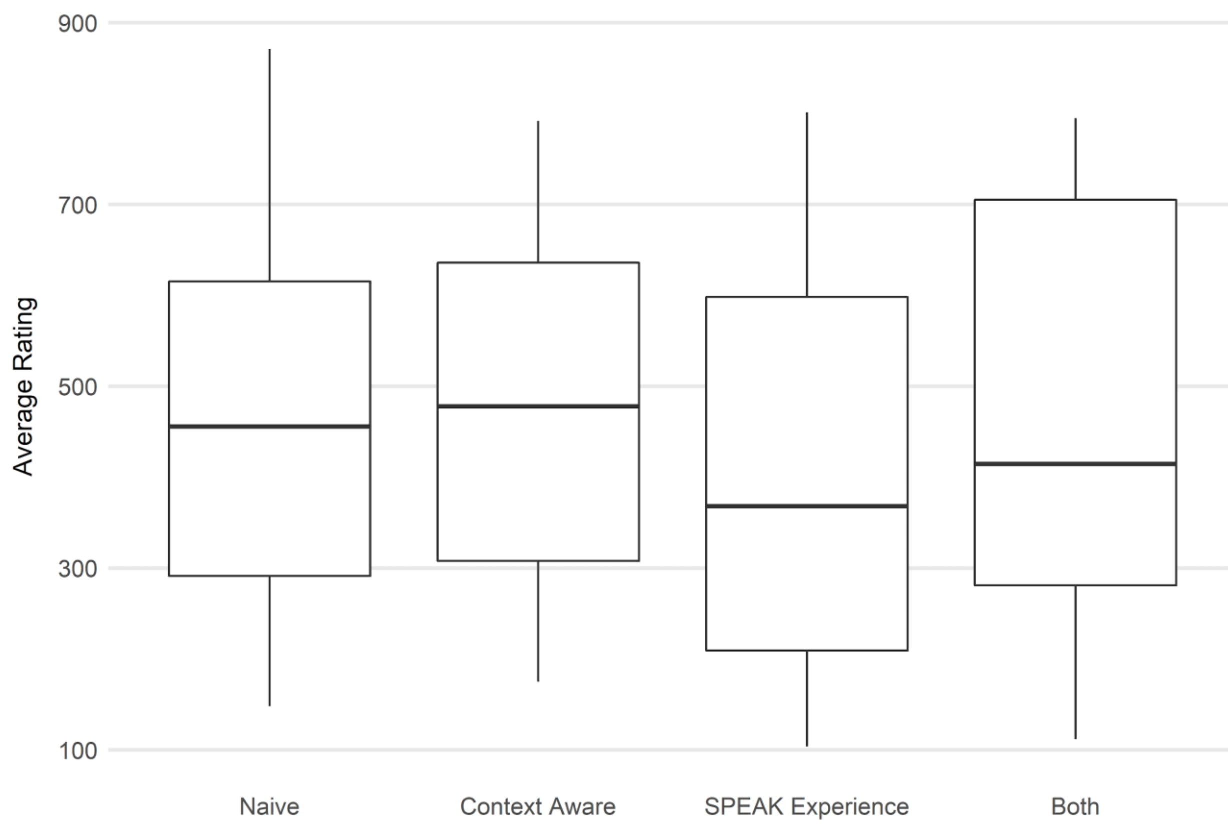
where $Y_{ijk}$ is the response, $\mu$ is the overall mean effect, and the terms $\alpha_i$, $\beta_j$, $(\alpha\beta)_{ij}$, and $\nu_{l(ij)}$ represent, respectively, the effect of the *i*th level of context awareness, the effect of the *j*th level of SPEAK experience, the effect of the interaction between $\alpha_i$ and $\beta_j$ (CA and SE), and the whole plot error (other sources of variability *between* the listener groups). Similarly, the terms $\gamma_k$, $(\alpha\gamma)_{ik}$, $(\beta\gamma)_{jk}$, and $\varepsilon_{l(ijk)}$ represent the effect of the *k*th level of proficiency, the effect of the interaction between $\alpha_i$ and $\gamma_k$ (CA and Prof), the effect of the interaction between $\beta_j$ and $\gamma_k$ (SE and Prof), and the subplot error (other sources of variability *within* the listener groups).

# 4. Results: Exploratory Data Analysis

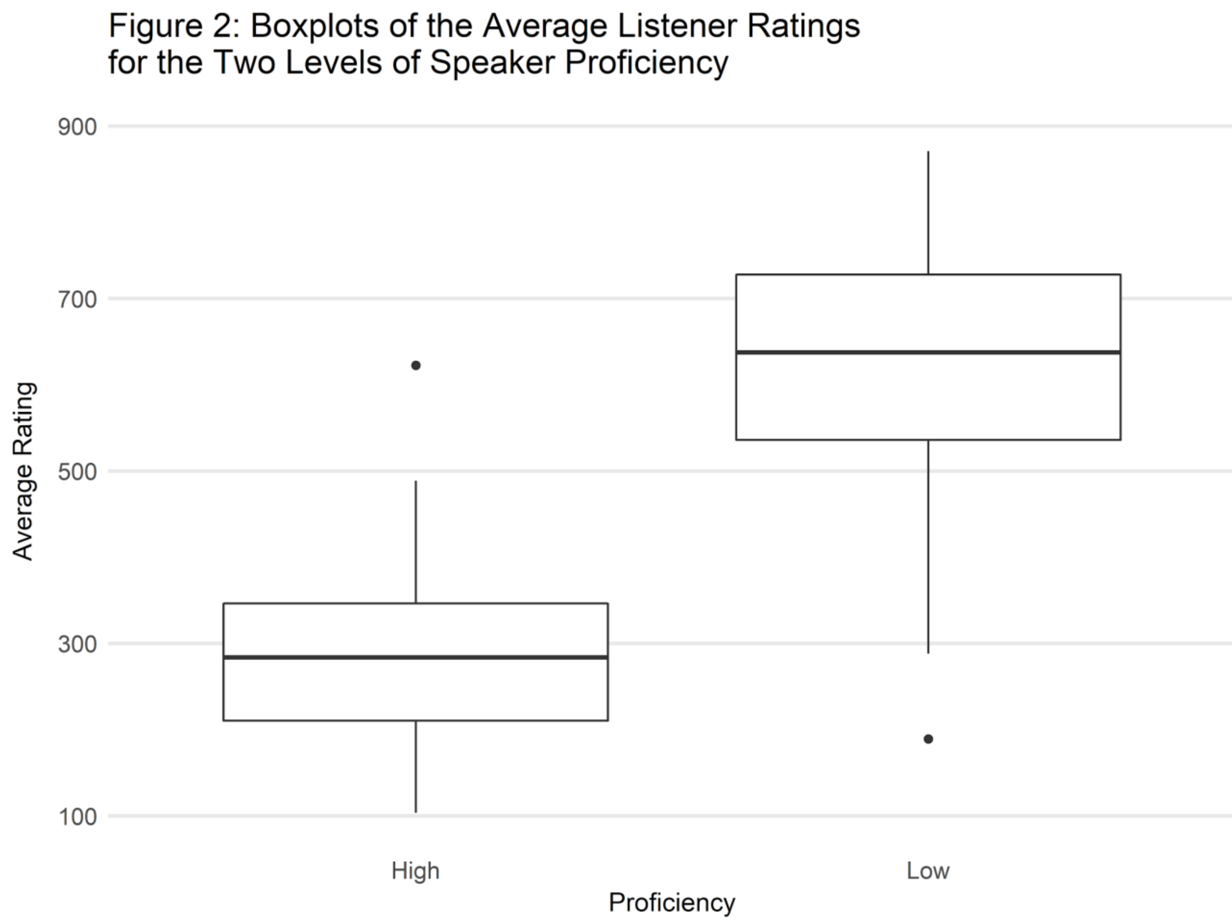
Before we began our analysis, we created some preliminary visualizations to explore the effects of the different variables of interest and the interactions between these variables.

For the following plots, the values on the vertical axis are the average ratings of accented speech (average of the Accentedness, Comfort, and Comprehensibility scores), where higher ratings correspond to lower perceptions of accented speech.

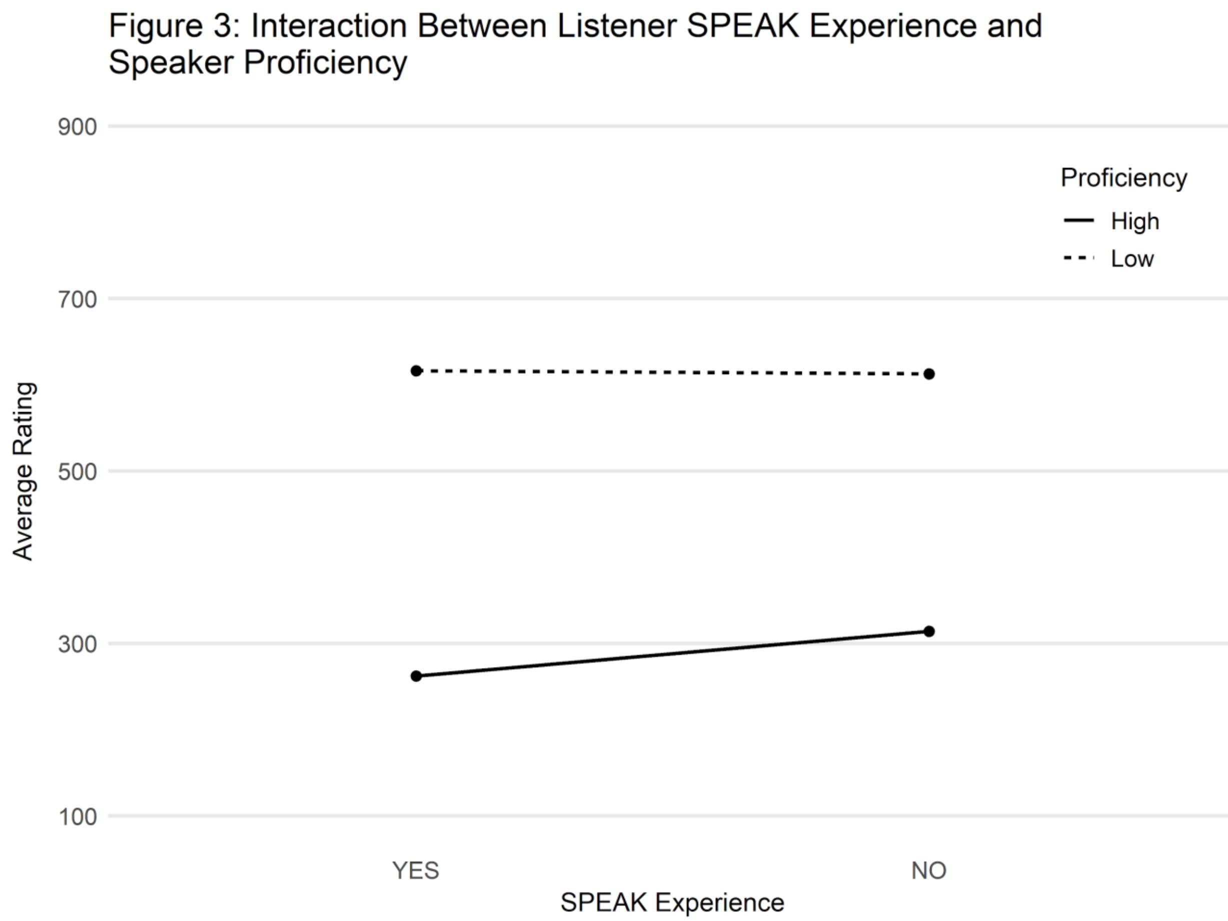
Figure 1: Boxplots of the Average Listener Ratings for the Four Listener Groups




Based on Figure 1, it does not appear that listeners who had SPEAK experience rated speakers differently than listeners who did not have SPEAK experience. Similarly, it does not appear that listeners who were context aware rated speakers differently than listeners who were not context aware.




Based on Figure 2, it appears that speakers were assigned different average ratings based on their level of proficiency. The low-proficiency speaker was scored much higher, on average, than the high-proficiency speaker.




Based on Figure 3, there may be an interaction between SPEAK experience and speaker proficiency. The lines are not parallel, which means the effect of SPEAK experience on average ratings is most likely different for different levels of speaker proficiency.

In addition, we also looked at plots of the interactions between context awareness and speaker proficiency and context awareness and SPEAK experience, but there did not seem to be interactions between them. The lines appeared parallel, which means the effect of one variable is likely the same regardless of the level of the other. These graphs are not included in this report.

# 5. Results: Modeling

Based on Table A.1 in the Modeling subsection of the Appendix, it appears that there is strong evidence (p-value < 0.001) to suggest that some of the variability in the average scores is due to differences in speaker proficiency. It also appears that there is some evidence (p-value = 0.0723) to suggest that some of the variability in the average scores is due to the interaction between SPEAK experience and speaker proficiency (SE*prof). We can see this in Figure 3 in the previous section: it appears that SPEAK experience has an effect on the average scores at the high level of proficiency, but not at the low level of proficiency. There is little evidence to suggest that some of the variability in the average scores is due to any of the other factors (p-values are extremely large). These results answer research questions 1-3 in the Introduction section.

Also, based on Table 5.1 below, there is little evidence to suggest that there is a difference in the effects of context awareness and test-taking experience on the averages of the accentedness, comprehensibility, and comfort scores (p-value = 0.2130), assuming the interaction between context awareness and SPEAK experience is 0. Table 5.1 looks at the difference between the estimated coefficients for context awareness and SPEAK experience from our model and performs a t-test. The null hypothesis for the test is that the difference between the estimated coefficients is statistically 0, and the alternative hypothesis is that the difference is statistically different from 0. This answers research question 4.

**Table 5.1:** t-test of the difference between the effects of CA and SE

| | |
|---|---|
| Value | -27.10565476 |
| Std Error | 21.499536741 |
| t Ratio | -1.260755294 |
| DF Den | 52 |
| Prob > \|t\| | 0.2130246836 |

To answer research question 5, we looked at the pairwise differences of context awareness and SPEAK experience interactions (CA-YES, SE-YES versus CA-NO, SE-YES and CA-YES, SE-NO) using the Tukey method of multiple comparisons and found that there is little evidence (Table A.2: p-values=0.4847, 0.9981) to suggest that there is a difference between the estimated average score of the “Both” listener group and either the “CA” or “SE” group.

# 6. Conclusions

## 6.1 Answers to Research Questions

1. How does experience taking a test of spoken English proficiency affect listener perceptions of accented speech?
    - There is little evidence (Table A.1: p-value = 0.4313) to suggest that experience taking a test of spoken English proficiency affects listener perceptions of accented speech.
2. How does context awareness affect listener perceptions of accented speech?

        - There is little evidence (Table A.1: p-value = 0.3269) to suggest that context awareness affects listener perceptions of accented speech.
3. Is there an interaction between listener condition (Naïve, CA, SE, Both) and speaker proficiency?
    - There is some evidence (Table A.1: p-value = 0.0723) to suggest that there is an interaction between experience taking a test of spoken English proficiency and speaker proficiency.
        - The differences between the combinations of this interaction only seemed important when the levels of proficiency were different (Table A.2: p-values < 0.0001).
        - This interaction explains less than 0.5% of the variation in the average scores within the listener groups; its contribution is negligible.
    - There is little evidence (Table A.1: p-value = 0.3044) to suggest that there is an interaction between context awareness and speaker proficiency.
4. Is there a significant difference in the effects of context awareness and test-taking experience on listener perceptions of accented speech? We understood this to mean a comparison between the sizes of the main effects for context awareness and SPEAK experience.
    - There is little evidence (Table A.3: p-value = 0.2130) to suggest that there is a difference between the effects of context awareness and test-taking experience on listener perceptions of accented speech.
5. Does the combination of experience taking a test of spoken English proficiency and context awareness result in different ratings than either test-taking experience or context awareness alone?
    - It does not appear that the combination of experience taking a test of spoken English proficiency and context awareness results in different ratings than either test-taking experience or context awareness alone; there is little evidence (Table A.2: p-values range from 0.4847 to 1.000) to suggest that any of the combinations of SPEAK experience and context awareness differ in the effects they have on listener perceptions of accented speech.

## 6.2 Caveats

Since the study only involved native Chinese speakers, we do not know the effect of different accents on the results.

Furthermore, we know from looking at the data that, of the 56 listeners, 54 of them were female and 36 (two-thirds) of these female participants were Speech Pathology & Audiology majors. As a result, we caution against generalizing beyond the population of female Miami University students in disciplines like Speech Pathology & Audiology.

# 7. References

Cooley, C. R. (2019). How do context awareness and experience taking the SPEAK test influence perceptions of non-native speaking proficiency? [PowerPoint Slides].

Montgomery, D. (2017). *Design and analysis of experiments* (Ninth ed.). Hoboken, NJ: John Wiley & Sons.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse.

Wickham, H., Byran, J. (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl.

# 8. Appendix

## 8.1 Modeling

**Table A.1:** SAS ANOVA table

| Type 3 Analysis of Variance | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **Expected Mean Square** | **Error Term** | **Error DF** | **F Value** | **Pr > F** |
| **CA** | 1 | 25359 | 25359 | Var(Residual) + 2 Var(listener(CA*SE)) + Q(CA,CA*SE,CA*prof) | MS(listener(CA*SE)) | 52 | 0.98 | 0.3269 |
| **SE** | 1 | 16286 | 16286 | Var(Residual) + 2 Var(listener(CA*SE)) + Q(SE,CA*SE,SE*prof) | MS(listener(CA*SE)) | 52 | 0.63 | 0.4313 |
| **prof** | 1 | 2979968 | 2979968 | Var(Residual) + Q(prof,CA*prof,SE*prof) | MS(Residual) | 53 | 464.36 | <.0001 |
| **CA*SE** | 1 | 27855 | 27855 | Var(Residual) + 2 Var(listener(CA*SE)) + Q(CA*SE) | MS(listener(CA*SE)) | 52 | 1.08 | 0.3044 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **CA*prof** | 1 | 70.634948 | 70.634948 | Var(Residual) + Q(CA*prof) | MS(Residual) | 53 | 0.01 | 0.9168 |
| **SE*prof** | 1 | 21586 | 21586 | Var(Residual) + Q(SE*prof) | MS(Residual) | 53 | 3.36 | 0.0723 |
| **listener(CA*SE)** | 52 | 1346014 | 25885 | Var(Residual) + 2 Var(listener(CA*SE)) | MS(Residual) | 53 | 4.03 | <.0001 |
| **Residual** | 53 | 340120 | 6417.365688 | Var(Residual) | . | . | . | . |

**Table A.2:** SAS Multiple Comparisons Table

| **Differences of Least Squares Means** | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Effect** | **CA** | **SE** | **prof** | **_CA** | **_SE** | **_prof** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > \|t\|** | **Adjustment** | **Adj P** | **Alpha** |
| **CA** | NO | | | YES | | | -30.0942 | 30.4049 | 52 | -0.99 | 0.3269 | Tukey | 0.3269 | 0.05 |
| **SE** | | NO | | | YES | | 24.1171 | 30.4049 | 52 | 0.79 | 0.4313 | Tukey | 0.4313 | 0.05 |
| **prof** | | | High | | | Low | -326.23 | 15.1391 | 53 | -21.55 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **CA*SE** | NO | NO | | NO | YES | | 55.6577 | 42.9991 | 52 | 1.29 | 0.2012 | Tukey | 0.5706 | 0.05 |
| **CA*SE** | NO | NO | | YES | NO | | 1.4464 | 42.9991 | 52 | 0.03 | 0.9733 | Tukey | 1.0000 | 0.05 |
| **CA*SE** | NO | NO | | YES | YES | | -5.9772 | 42.9991 | 52 | -0.14 | 0.8900 | Tukey | 0.9990 | 0.05 |
| **CA*SE** | NO | YES | | YES | NO | | -54.2113 | 42.9991 | 52 | -1.26 | 0.2130 | Tukey | 0.5917 | 0.05 |
| **CA*SE** | NO | YES | | YES | YES | | -61.6349 | 42.9991 | 52 | -1.43 | 0.1577 | Tukey | 0.4847 | 0.05 |
| **CA*SE** | YES | NO | | YES | YES | | -7.4236 | 42.9991 | 52 | -0.17 | 0.8636 | Tukey | 0.9981 | 0.05 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **CA*prof** | NO | | High | NO | | Low | -327.82 | 21.4099 | 53 | -15.31 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **CA*prof** | NO | | High | YES | | High | -31.6825 | 33.9654 | 53 | -0.93 | 0.3552 | Tukey-Kramer | 0.7874 | 0.05 |
| **CA*prof** | NO | | High | YES | | Low | -356.33 | 33.9654 | 53 | -10.49 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **CA*prof** | NO | | Low | YES | | High | 296.14 | 33.9654 | 53 | 8.72 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **CA*prof** | NO | | Low | YES | | Low | -28.5060 | 33.9654 | 53 | -0.84 | 0.4051 | Tukey-Kramer | 0.8355 | 0.05 |
| **CA*prof** | YES | | High | YES | | Low | -324.64 | 21.4099 | 53 | -15.16 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **SE*prof** | | NO | High | | NO | Low | -298.47 | 21.4099 | 53 | -13.94 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **SE*prof** | | NO | High | | YES | High | 51.8829 | 33.9654 | 53 | 1.53 | 0.1326 | Tukey-Kramer | 0.4286 | 0.05 |
| **SE*prof** | | NO | High | | YES | Low | -302.12 | 33.9654 | 53 | -8.89 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **SE*prof** | | NO | Low | | YES | High | 350.35 | 33.9654 | 53 | 10.31 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |
| **SE*prof** | | NO | Low | | YES | Low | -3.6488 | 33.9654 | 53 | -0.11 | 0.9149 | Tukey-Kramer | 0.9996 | 0.05 |
| **SE*prof** | | YES | High | | YES | Low | -354.00 | 21.4099 | 53 | -16.53 | <.0001 | Tukey-Kramer | <.0001 | 0.05 |

## 8.2 Data

**Original Data:** “SPEAK Results with Ciara’s data 2019_9_10 with comments.xlsx”; 12 columns and 1344 rows

| Variable Name | Description |
|---|---|
| Listener | The group code (N, CA, SE, B) followed by a unique number for each of the 56 listeners. |
| Sound | The file name; the first three characters correspond to the speaker and the rest to the speaking task. Each speaker has 12 associated audio files. |
| Description | Short description of the speaking task. |
| Proficiency | Proficiency of the speaker; “High” or “Low”. |
| Accentedness | Perceived strength of the speaker’s accent as rated by each listener for each audio file; integer value from 100-900 where higher values correspond to lower accentedness. |
| Comprehensibility | How easy or hard to understand the speaker as rated by each listener for each audio file; integer values from 100-900 where higher values correspond to lower comprehensibility. |
| Comfort | Comfort level with having this speaker as an instructor as rated by each listener for each audio file; integer values from 100-900 where higher values correspond to lower comfort. |
| SPEAK Experience | Whether the listener had taken a different form of the SPEAK test (“YES” or “NO”). |
| Context Aware | Whether the listener was aware of the topics of discourse (that is, the topic of each audio file); “YES” or “NO”. |
| Code | One of four groups that listeners were randomly assigned: N-naïve, CA - context aware, SE - SPEAK experience, B - have both context awareness and SPEAK experience. |
| SPA Major | Whether the listener was a speech pathology and audiology major (“YES” or “NO”). |
| Undergraduate | Whether the listener was an undergraduate student (“YES” or “NO”) |
| Age | Age of the listener; integer. |
| Gender | Gender of the listener; “M” for male and “F” for female. |

| **Altered Data**: “speech.csv”; 5 columns and 112 rows | |
|---|---|
| **Variable Name** | **Description** |
| listener | The group code (N, CA, SE, B) followed by a unique number for each of the 56 listeners. |
| CA | Whether the listener was aware of the topics of discourse (that is, the topic of each audio file); “YES” or “NO”. |
| SE | Whether the listener had taken a different form of the SPEAK test (“YES” or “NO”). |

| | |
|---|---|
| prof | Proficiency of the speaker; “High” or “Low”. |
| Avg_Score | Average of the 36 scores (12 audio files, each with an Accentedness, Comprehensibility, and Comfort score) from each listener for the two speaker proficiency levels. |

## 8.3 R Code

```
## Required packages
library(ggplot2)
library(readxl)
library(tidyverse)
## Read in the data
speech <- read_excel("C:/Users/tyler/Documents/STA 475/Client 2/SPEAK Results
with Ciara's data 2019_9_10 with comments.xlsx", col_names = TRUE)

## Change the variable names
names(speech) <- c("listener", "sound", "desc", "prof", "acc", "comp",
"comf", "SE", "CA", "code", "SPA", "UG", "age", "gender")

## Make comfort and comprehensibility numeric
speech <- speech %>%
mutate(comp = as.numeric(comp),
  comf = as.numeric(comf),
  CA = toupper(CA))

## Reordering the factor levels for some variables
speech$SE <- factor(speech$SE, c("YES", "NO"))

speech$CA <- factor(speech$CA, c("YES", "NO"))

speech$SPA <- factor(speech$SPA, c("YES", "NO")) ## Reorder the levels
levels(speech$SPA) <- c("SPA", "Other") ## Relabel the levels

speech$prof <- factor(speech$prof, c("High", "Low"))

speech$code <- factor(speech$code, c("N", "CA", "SE", "B"))
levels(speech$code) <- c("Naive", "Context Aware", "SPEAK Experience",
"Both")

## Average the accentedness, comfort, and comprehensibility scores
speech$percep <- (speech$comf + speech$comp + speech$acc)/3

## Condense the data into 5 columns and 112 rows
speech <- speech %>%
  group_by(listener, CA, SE, prof, code) %>%
  summarise(Avg_Score = mean(percep))

## Output the new data set as a .csv file for use in SAS
write.csv(speech, file="speech.csv")
## Custom theme
t1 <- theme(axis.text = element_text(size=11),
        axis.title.x = element_text(size=12, margin = margin(t=4, b=4)),
```

```
        axis.title.y = element_text(size=12, margin = margin(r=10)),
        legend.text = element_text(size=11),
        legend.title = element_text(size=12),
        legend.key = element_rect(fill="white", color="white"),
        legend.position = c(0.85,0.38),
        title = element_text(size=13),
        plot.caption = element_text(size=11, color="gray40", hjust = 0),
        panel.background = element_blank(),
        panel.border = element_blank(),
        axis.ticks = element_blank(),
        panel.grid.major.y = element_line(color="#e8e8e8", size=0.8),
        panel.grid.minor.y = element_blank(),
        panel.grid.major.x = element_blank())

## Used in all plots
s1 <- scale_y_continuous(name = "Average Rating",
                    breaks = c(100, 300, 500, 700, 900),
                    limits = c(100, 900))
p1 <- ggplot(data=speech) +
  geom_boxplot(aes(x=code, y=Avg_Score)) +
  labs(title="Figure 1: Boxplots of the Average Listener Ratings \nfor the
Four Listener Groups") +
  t1 +
  s1 +
  theme(axis.title.x = element_blank())

p1
p2 <- ggplot(data=speech) +
  geom_boxplot(aes(x=prof, y=Avg_Score)) +
  labs(x = "Proficiency", title="Figure 2: Boxplots of the Average Listener
Ratings \nfor the Two Levels of Speaker Proficiency") +
  t1 +
  s1

p2
## Group to get average scores for each combination of
## SPEAK experience and proficiency
speech2 <- speech %>%
  group_by(SE, prof) %>%
  summarise(Avg_Percep = mean(Avg_Score))

## Interaction Plot of SPEAK Experience and Proficiency
p3 <- ggplot(data=speech2) +
  geom_point(aes(x=SE, y=Avg_Percep), size=2) +
  geom_line(aes(x=SE, y=Avg_Percep, group=prof, lty=prof),
          color= "black", size=0.8) +
  scale_linetype_discrete(name = "Proficiency", limits=c("High", "Low")) +
  s1 +
  labs(x="SPEAK Experience", y="Average Rating", title="Figure 3: Interaction
Between Listener SPEAK Experience and \nSpeaker Proficiency") +
  t1

p3
ggsave("C:/Users/tyler/Documents/STA 475/Client 2/group_plot.png", plot=p1,
width=8, height=6, units="in", dpi=600)
ggsave("C:/Users/tyler/Documents/STA 475/Client 2/prof_plot.png", plot=p2,
width=8, height=6, units="in", dpi=600)
```

```
ggsave("C:/Users/tyler/Documents/STA 475/Client 2/SE_prof_plot.png", plot=p3,
width=8, height=6, units="in", dpi=600)
```

## 8.4 SAS Code

```
%let work_dir = \\tsclient\C\Users\tyler\Documents\STA 475;

proc import datafile= "&work_dir\speech.csv" out=speech dbms=csv replace;
      getnames=yes;
run;

data speech;
      set speech;
      drop VAR1;
run;

proc mixed NOINFO method=TYPE3 data=speech;
      class CA SE prof listener;
      model Avg_Score=CA SE prof CA*SE CA*prof SE*prof / outp=ps;
      lsmeans CA SE prof CA*SE CA*prof SE*prof / adjust=Tukey cl;
      random listener(SE*CA);
run;
```