COGS 219 R Notebook Assignment Code ▾

An Event-Related Potential Study of Music and Language Processing

Rena A. Lee

COGS-219-01

Spring 2018

Introduction

The present study was an attempt at replication of Patel et al.'s *Processing Syntactic Relations in Language and Music: An Event-Related Potential Study* published in 1998. Patel et al. used event-related potentials to investigate language specificity in neural syntactic processing, directly comparing the P600 ERP wave components produced by syntactic incongruities in language and in music. In the original study, stimuli for both language and music were constructed in three conditions: congruous, moderately incongruous, and highly incongruous. Fifteen musically experienced adults participated in the within-subjects designed experiment, and it was concluded that linguistic and musical syntactic anomalies produced the same ERP component in a specified latency range (450-750 ms). Additionally, evidence of right-hemisphere lateralization of music processing was found. In all, their results support the notion that neural resources for processing syntax are not language-specific, and more broadly give way for investigating neural specificity in cognitive processing.

The purpose of conducting this replication was to become more familiar with themes of replicability, reproducibility, and open science in the context of cognitive science research. In the broad scope of cognitive science research (and research in related fields such as psychology), there has been a replication crisis, where it has recently been discovered that a large proportion of published articles are unable to be corroborated by subsequent testing. Another issue is that failed replications or original studies with null results often go unreported. This is an issue because, as Nosek (2017) puts it, "Open science is the key to reducing waste, accelerating meaningful solutions to the biggest problems faced by our communities, states, nations, business and civic institutions, and to save the lives of millions of people around the world…" There is a stark disconnect between the scientific community and the general public, which desperately needs to be remedied.

To contribute to solving widespread issues of replicability and open science, we preregistered this study on the Open Science Framework and made supplemental materials, methods, and other information openly available to those who seek it. Power (2016) asserts that registered reports should be the standard among scientific practices in order for truthful information to be made available to the public, so that others may attempt to replicate your results as well. Having been in contact with the original head author personally, we were able to reproduce the experiment without much difficulty. Certain aspects of the original study were changed for our methods, but overall the study was designed to be a very close reproduction. The aim was to replicate Patel et al.'s results which supported the notion on language-music shared processing features in the brain.

Methods

The following is a summary of the methods used in this replication. The full details of our methods can be found in our preregistration at https://osf.io/g3b5j/ (https://osf.io/g3b5j/). The stimuli, experiment scripts (using jsPsych on JavaScript), data, and analysis files utililzed may be accessed in our OSF repository: https://osf.io/zpm9t/ (https://osf.io/zpm9t/).

Subjects with at least four years of broadly concieved musical experience were sampled from the Vassar College student body; with a target subject pool of n=40, a total of 44 subjects were recruited through email, class announcements, social media, and flyers; 35 were included in our analyses.

The stimuli used for this replication were virtually the same as Patel et al.'s. Language stimuli included grammatical and ungrammatical syntactically structured sentences; 100 total sentences were randomly assigned into three blocks of 33, 33, and 34 sentences. A female researcher voice recorded the sentences, which were provided by Patel. The original condition of 'moderately incongruous,' for both language and music stimuli, were not included in this replication. Music stimuli included in-key and distant-key chords within short musical phrases; there were two blocks each containing 36 music stimuli. These audio files were directly recieved from Patel.

Participants were fitted with an EGI net according to their experimenter-measured head circumference. Once the net was applied, they were asked to complete the experimental task: identifying whether sentences and musical phrases were acceptable or unnaceptable based on syntax and chord structure, respectively. During randomized controlled trials of alternating sentence and music samples (according to a within-subjects block style design with no counterbalancing), EEG data was recorded from 100 ms before to 1000 ms after target stimulus onset for each trial. Basing off the Patel et al. study, 13 of the 128 electrodes on the net were used in the analysis: 33, 39, 45, 70, 11, 83, 62, 122, 115, 108, 42, 93, and REF (Cz). Once the task was completed, participants were disconnected from the system, the EGI net was removed, and they were debriefed about the background of the experiment.

Once the raw EEG data was collected, it was transformed using NetStation EGI software using the following tools: 0.10-30.00 Hz bandpass filter, bad channel replacement, average re-referencing, and baseline correction. After this pre-processing phase, segmented data for each trial was manually reviewed for anomalies and then recorded by student experimenters into a spreadsheet to collect minimum number of good segments per stimulus category for each participant. A total of eight statistical tests were used in data analysis: six classic repeated-measures ANOVAs and two Bayesian model repeated-measures ANOVAs, reported and discussed in the EEG Data Analysis section below.

Results

Demographic Data Analysis

The demographic data collected includes gender identity, date of birth, the instrument(s) they have experience with, estimated years of total musical experience, and hours spent per week involving musical practice. Subjects filled out surveys to provide this information before experiments were run.

- **Subject** is the Participant ID. n=35
- **Gender** is either F (female), M (male), or N (non-binary).
- **YearsMusicalExperience** is the participant's musical experience, regardless of how many instruments they may have played.
- **HoursPerWeek** is the participant's self-estimated time spent engaged in musical practice per week.

Demographic Data Summary

Below is a table of the demographic data collected, excluding subjects 8, 10, 11, 21, 23, 25, and 40. Due to experimenter error and/or participant exclusionary criteria (had perfect pitch, impossible net application, failure to show), subjects 8, 10, 21, 27, and 30 were excluded from this analysis as they did not produce any EEG data. The data produced by subjects 11, 23, 25, and 40 were excluded from analysis because they did not produce the minimum number of good segments per stimulus category in their EEG data, which was 19. Segments were determined to be bad by NetStation artifact (eye blinks/movements) detection software. Said subjects were filtered out the the raw demographic data data. The data was also edited to read Subject as a factor rather than an integer value.

Hide

```
demographic_data <- read.csv('demographics.csv')
demographic_data$Subject <- factor(demographic_data$Subject)
filtered_demographic_data <- demographic_data %>% filter(!Subject %in% c(8, 11, 21, 23,
25, 40))
filtered_demographic_data
```

Subject	Gen...	Birthdate	Instruments	YearsMusicalExperience	HoursPerWe
<fctr>	<fctr>	<fctr>	<fctr>	<int>	<in
1	F	03/19/1998	piano, flute	6	
2	M	01/11/1997	piano, drums	12	
3	M	01/15/1997	guitar, violin	8	
4	F	04/12/1998	vioin, voice	6	
5	M	01/21/1997	guitar, piano	5	
6	F	03/19/1999	flute	11	
7	M	09/03/1996	oboe	4	
9	F	01/22/1999	clarinet	7	
12	M	10/21/1996	cello, guitar, piano	16	
13	F	09/29/1998	violin	15	

1-10 of 35 rows   Previous   **1**   2   3   4   Next

Age Demographic

Participant ages at the time of experimentation were extrapolated from their birthdates.

Hide

```
date_now = as.Date("02/25/2018", format = "%m/%d/%Y")
birthday_demographic <- filtered_demographic_data %>%
  mutate(Birthdate = as.character(Birthdate))%>%
  mutate(Birthdate = as.Date(Birthdate,format = "%m/%d/%Y")) %>%
  mutate(age_years = age_calc(Birthdate,date_now,units = "years"))
```

Means of Demographic Data

Below is a table summarizing the mean age, years of musical experience, and hours spent engaged in musical activities per week of the participants of our study. Respective standard deviations are displayed to the right of each mean value.

Hide

```
Means <- birthday_demographic %>%
  summarize(MeanAge = round(mean(age_years), 2),
            SDAge = round(sd(age_years), 2),
            MeanYearsExperience = round(mean(YearsMusicalExperience), 2),
            SDYearsExperience = round(sd(YearsMusicalExperience), 2),
            MeanHoursPerWeek = round(mean(HoursPerWeek), 2),
            SDHoursPerWeek = round(sd(HoursPerWeek), 2))
Means
```

Mean...	SD...	MeanYearsExperience	SDYearsExperience	MeanHoursPerW...	SDHoursPerW...
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
20.32	1.12	9.71	3.43	5.83	3.42

1 row

Gender Frequencies

There were 20 female, 14 male, and 1 non-binary participants in the study.

Behavioral Data Analysis

Behavioral data includes the following:

- **Subject** is the Participant ID. As discussed in the Demographic Data section, participants 8, 10, 11, 21, 23, 25, and 40 were excluded from this analysis. n=35.
- **RT** is the reaction time, in milliseconds, of the participant for each trial.
- **KeyPressed** is the key that the participant pressed in response to the stimulus. U = Ungrammatical, A = Grammatical
- **StimulusType** is either Sentence or Music.
- **SyntaxCategory** is either Ungrammatical or Grammatical for the Sentence stimuli, and either Distant-Key or In-Key for the Music stimuli. Filler sentences were not included in this analysis.
- **Correct** indicates whether the participant correctly identified the stimulus as acceptable or unacceptable.

Below is a table summarizing the behavioral data of each participant. Data from excluded participants and filler condition stimuli were filtered out of the raw behavioral data from the Subject and SyntaxCategory columns, respectively.

Hide

```
behavioral_data <- read.csv('beh_data_tidy.csv')
behavioral_data$Subject <- factor(behavioral_data$Subject)
behavioral_data$KeyPressed[behavioral_data$KeyPressed == '85'] <- "U"
behavioral_data$KeyPressed[behavioral_data$KeyPressed == '65'] <- "A"
filtered_behavioral_data <- behavioral_data %>% filter(!Subject %in% c(8, 10, 11, 21, 23
, 25, 27, 30, 40)) %>%
  filter(!SyntaxCategory %in% c('Filler-Gram', 'Filler-Ungram'))
filtered_behavioral_data
```

Subject	RT	KeyPressed	StimulusType	SyntaxCategory	Correct
<fctr>	<int>	<chr>	<fctr>	<fctr>	<lgl>
1	1766	U	Sentence	Ungrammatical	TRUE
1	1983	U	Sentence	Ungrammatical	TRUE
1	1774	U	Sentence	Ungrammatical	TRUE
1	1674	A	Sentence	Grammatical	TRUE
1	1388	A	Sentence	Ungrammatical	FALSE
1	1784	A	Sentence	Grammatical	TRUE
1	1744	A	Sentence	Ungrammatical	FALSE
1	6588	A	Sentence	Grammatical	TRUE
1	725	A	Sentence	Grammatical	TRUE
1	1506	U	Sentence	Ungrammatical	TRUE

1-10 of 4,620 rows   Previous **1** 2 3 4 5 6 … 100 Next

Accuracy Data

The proportion of correct responses to total responses for each participant was calculated, separated into the four critical conditions of stimulus type and syntax category. Patel referred to this as "percent endorsement." The proportion of correct responses was calculated first for each participant, and then the mean of the proportion of correct responses among all participants, along with the standard deviation, was calculated and is displayed below.

Hide

```
correct_response_data <- filtered_behavioral_data %>%
  group_by(StimulusType, SyntaxCategory, Subject) %>%
  summarize(ProportionCorrect = (mean(Correct))) %>%
  group_by(StimulusType, SyntaxCategory) %>%
  summarize(MeanProportionCorrect = mean(ProportionCorrect), SD = sd(ProportionCorrect))
correct_response_data
```

StimulusType	SyntaxCategory	MeanProportionCorrect	SD
<fctr>	<fctr>	<dbl>	<dbl>

StimulusType	SyntaxCategory	MeanProportionCorrect	SD
<fctr>	<fctr>	<dbl>	<dbl>
Music	Distant-Key	0.6928571	0.15727095
Music	In-Key	0.8428571	0.14666822
Sentence	Grammatical	0.9342857	0.05392182
Sentence	Ungrammatical	0.8819048	0.18245652
4 rows			

EEG Data Analysis

EEG data collected from the participants whom were excluded due to reasons discussed above (see Demographic Data Analysis) has been filtered out. Additionally, variables were renamed to maintain consistency with the previous two data files. (This was a purely stylistic change.)

Hide

```
EEG_data <- read.csv('eeg_data_tidy.csv')
EEG_data$Subject <- factor(EEG_data$subject)
EEG_data$subject <- NULL
EEG_data$Electrode <- factor(EEG_data$electrode)
EEG_data$electrode <- NULL
EEG_data$Time <- EEG_data$t
EEG_data$t <- NULL
EEG_data$Voltage <- EEG_data$voltage
EEG_data$voltage <- NULL
EEG_data$StimulusCondition <- EEG_data$stimulus.condition
EEG_data$stimulus.condition <- NULL
EEG_data$GrammarCondition <- EEG_data$grammar.condition
EEG_data$grammar.condition <- NULL
filtered_EEG_data <- EEG_data %>% filter(!Subject %in% c(11, 23, 25, 40))
```
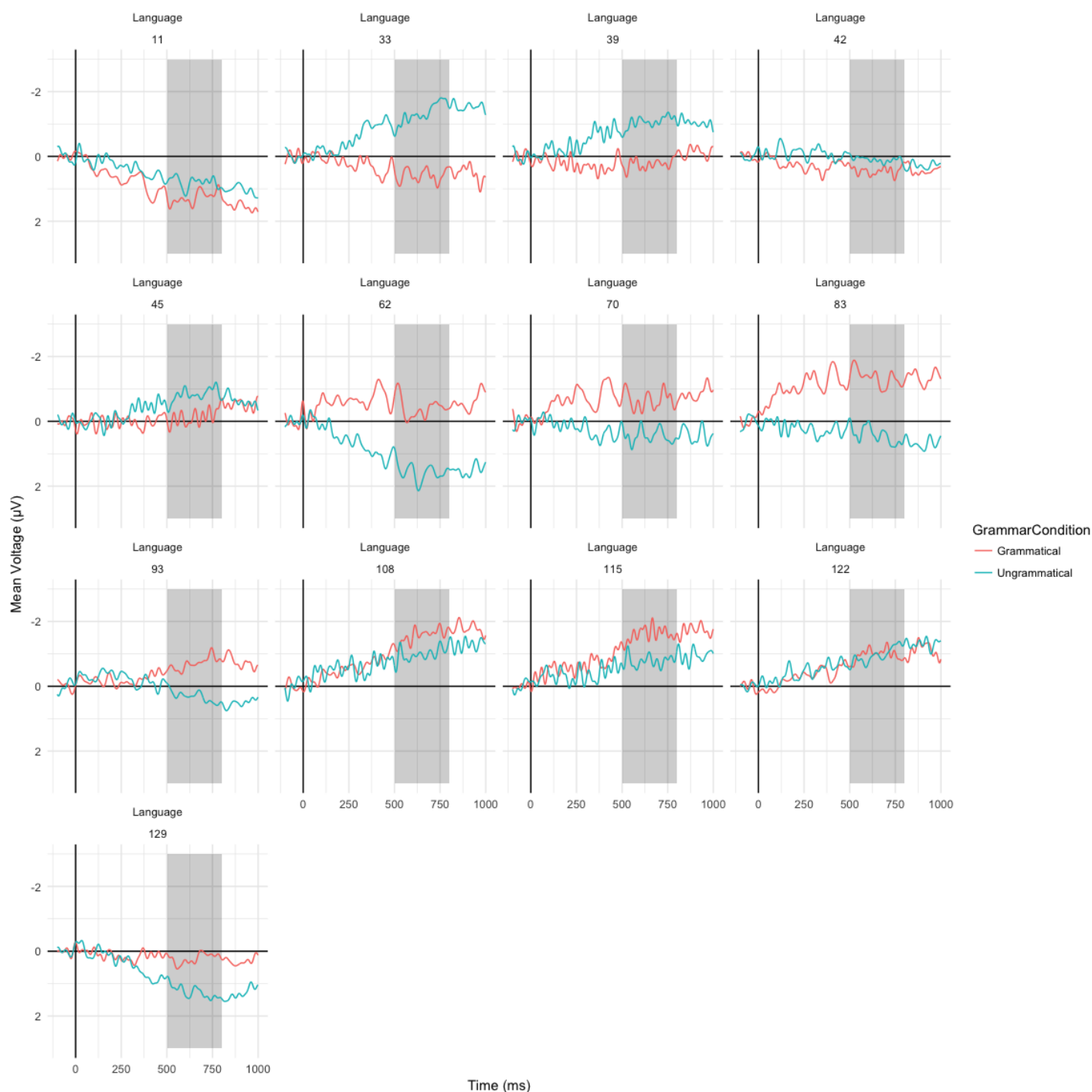
Grand Averages

The figures below display the mean voltages of each of the electrodes of interest (n=13) separated by stimulus condition. The highlighted section of each plot is the time window of interest, 500 ms to 800 ms after stimulus onset.

Language Grand Averages Plot

The filtered EEG data was further filtered to include average waveforms for only the language stimulus condition. The mean voltages across subjects for each electrode was calculated, and plotted on the graph below.

Hide

```
grand_average_dataL <- filtered_EEG_data %>%
    filter(StimulusCondition %in% c('Language')) %>%
    group_by(Time, Electrode, StimulusCondition, GrammarCondition) %>%
    summarize(MeanVoltage = mean(Voltage))
ggplot(grand_average_dataL, aes(x=Time, y=MeanVoltage, color = GrammarCondition)) +
    annotate("rect", xmin=500, xmax=800, ymin=-3, ymax=3, fill="black", alpha=0.2) +
    geom_vline(xintercept = 0) +
    geom_hline(yintercept = 0) +
    geom_line() +
    facet_wrap(StimulusCondition ~ Electrode) +
    labs(x='Time (ms)', y='Mean Voltage (µV)') +
    scale_y_reverse() +
    theme_minimal()
```

Music Grand Averages Plot

The filtered EEG data was further filtered to include average waveforms for only the music stimulus condition. The mean voltages across subjects for each electrode was calculated, and plotted on the graph below.

Hide

```
grand_average_dataM <- filtered_EEG_data %>%
  filter(StimulusCondition %in% c('Music')) %>%
  group_by(Time, Electrode, StimulusCondition, GrammarCondition) %>%
  summarize(MeanVoltage = mean(Voltage))
ggplot(grand_average_dataM, aes(x=Time, y=MeanVoltage, color = GrammarCondition)) +
  annotate("rect", xmin=500, xmax=800, ymin=-3, ymax=3, fill="black", alpha=0.2) +
  geom_vline(xintercept = 0) +
  geom_hline(yintercept = 0) +
  geom_line() +
  facet_wrap(StimulusCondition ~ Electrode) +
  labs(x='Time (ms)', y='Mean Voltage (µV)') +
  scale_y_reverse() +
  theme_minimal()
```
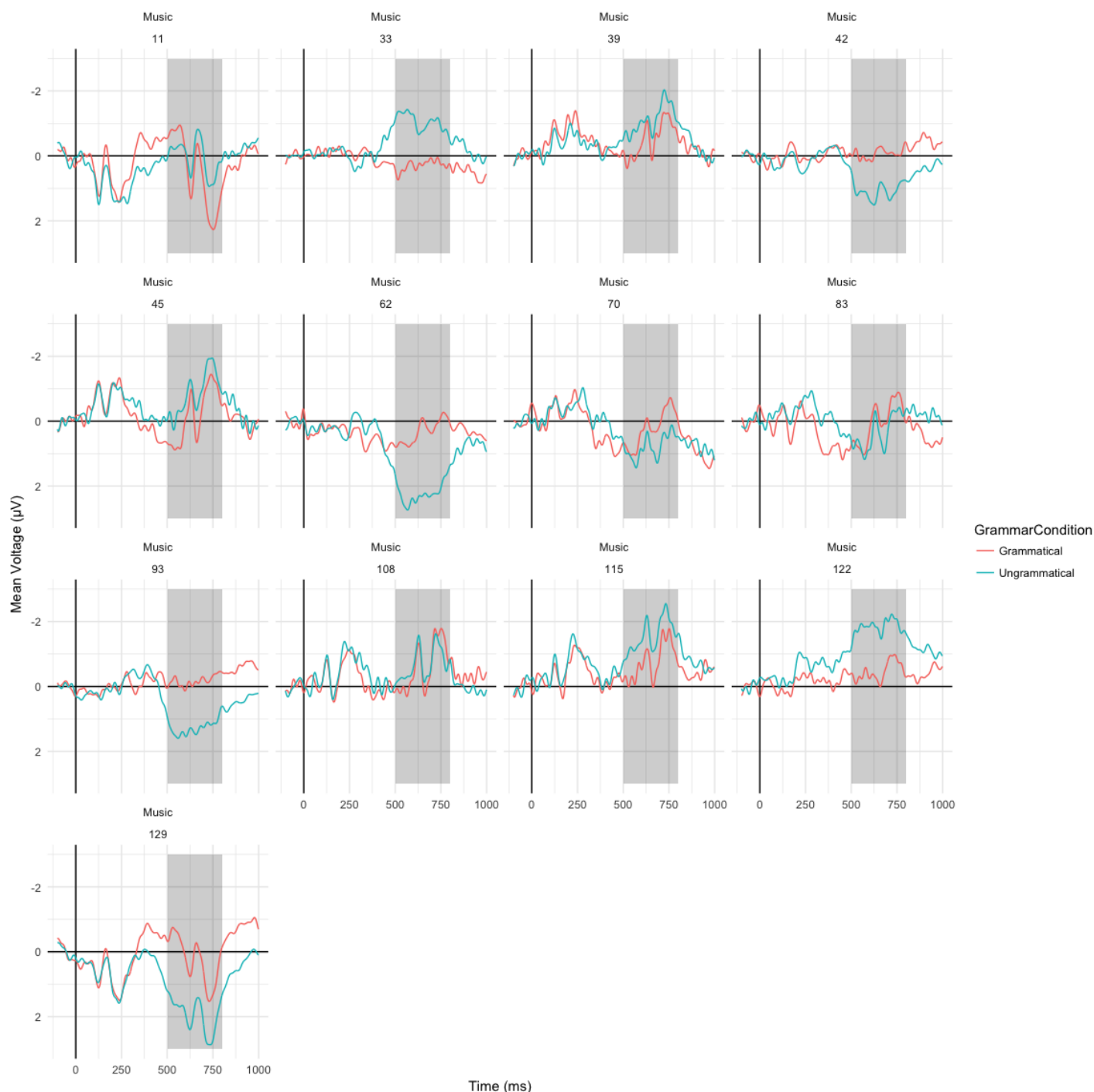
Statistical Analyses of EEG Data

Following Patel et al.'s analyses, six repeated-measures ANOVAs were run to observe the possible effect of stimulus type and grammar condition on ERP waveforms. Our anovas aren't directly comparable to theirs because they had three grammar conditions and we had two, but the same effects and relationships can be investigated with sufficient power. The first four ANOVAs are investigating the effect of grammar condition (grammatical or ungrammatical) on the averaged ERP amplitudes. The alpha level was set to 0.05. Our registration outlines this process as follows:

> Repeated-measures ANOVA, 2 (grammatical v. nongrammatical) x N (where N = the number of relevant electrodes; see below), of the mean amplitude of wave forms will be conducted in the 500 to 800 msec window separately for both the language and music conditions and separately for lateral and midline electrodes, following Patel et al.

ANOVA: Midline Electrodes, Language

Below is a 2 by 3 factor repeated-measures ANOVA of the mean voltage for each electrode produced by each participant for language stimuli. EEG data was filtered to the time window of interest, 500 to 800 ms, and the three midline electrodes, 11 (Fz), 62 (Pz), and 129 (CzRef). It appears that the greater mean voltages in the ungrammatical condition in electrodes 62 and 129 are responsible for the effect of grammar condition.

Hide

```
ANOVA_EEG_dataM <- filtered_EEG_data %>%
  filter(Time %in% 500:800, Electrode %in% c(11, 62, 129)) %>%
  group_by(Subject, Electrode, GrammarCondition, StimulusCondition) %>%
  summarize(MeanVoltage = mean(Voltage))
EEG_ANOVA_resultML <- ezANOVA((ANOVA_EEG_dataM %>% filter(StimulusCondition == 'Languag
e')), dv = MeanVoltage, wid=Subject, within=c(GrammarCondition, Electrode))
EEG_ANOVA_resultML$ANOVA
```

Effect	...	...	F	p	p<.05	ges
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>
2 GrammarCondition	1	34	6.412910	0.016114742	*	0.036350095
3 Electrode	2	68	1.003302	0.372028824		0.007251095
4 GrammarCondition:Electrode	2	68	5.115345	0.008520782	*	0.050927978

3 rows

Hide

```
ML_descriptives <- ANOVA_EEG_dataM %>%
  filter(StimulusCondition %in% 'Language') %>%
  group_by(Electrode, GrammarCondition) %>%
  summarize(Mean = mean(MeanVoltage), SD = sd(MeanVoltage))
ML_descriptives
```

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
11	Grammatical	1.2668727	1.720958
11	Ungrammatical	0.7946868	2.440684
62	Grammatical	-0.4100343	3.030211

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
62	Ungrammatical	1.5710685	2.500408
129	Grammatical	0.2252987	1.481970
129	Ungrammatical	1.2569022	1.675383
6 rows			

ANOVA: Midline Electrodes, Music

Below is a 2 by 3 factor repeated-measures ANOVA of the mean voltage produced by each participant for music stimuli. EEG data was filtered to the time window of interest, 500 to 800 ms, and the three midline electrodes, 11, 62, and 129. Similarly to the above ANOVA, electrodes 62 and 129 elicited significantly greater mean voltages for the distant-key condition compared to the in-key condition, accounting for the effect of grammar condition on the grand average waveforms produced in the midline electrodes.

Hide

```
EEG_ANOVA_resultMM <- ezANOVA((ANOVA_EEG_dataM %>% filter(StimulusCondition == 'Music'
)), dv = MeanVoltage, wid=Subject, within=c(GrammarCondition, Electrode))
EEG_ANOVA_resultMM$ANOVA
```

Effect	...	...	F	p	p<.05	ges
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>
2 GrammarCondition	1	34	23.93652	2.366626e-05	*	0.09596085
3 Electrode	2	68	6.99996	1.720468e-03	*	0.06719087
4 GrammarCondition:Electrode	2	68	12.43002	2.506803e-05	*	0.09589917
3 rows						

Hide

```
MM_descriptives <- ANOVA_EEG_dataM %>%
  filter(!StimulusCondition %in% 'Language') %>%
  group_by(Electrode, GrammarCondition) %>%
  summarize(Mean = mean(MeanVoltage), SD = sd(MeanVoltage))
MM_descriptives
```

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
11	Grammatical	0.48891625	1.564918
11	Ungrammatical	0.05797916	1.514645
62	Grammatical	0.31185122	1.424570
62	Ungrammatical	2.22239531	2.511751

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
129	Grammatical	0.25822741	1.014809
129	Ungrammatical	1.94439780	1.444951
6 rows			

ANOVA: Lateral Electrodes, Language

Below is a 2 by 10 factor repeated-measures ANOVA of the mean voltage produced by each participant for language stimuli. EEG data was filtered to the time window of interest, 500 to 800 ms, and the ten lateral electrodes, 33, 39, 42, 45, 70, 83, 93, 108, 115, and 122. There is no observed effect of grammar condition.

Hide

```
ANOVA_EEG_dataL <- filtered_EEG_data %>%
  filter(Time %in% 500:800, !Electrode %in% c(11, 62, 129)) %>%
  group_by(Subject, Electrode, GrammarCondition, StimulusCondition) %>%
  summarize(MeanVoltage = mean(Voltage))
EEG_ANOVA_resultLL <- ezANOVA((ANOVA_EEG_dataL %>% filter(StimulusCondition == 'Languag
e')), dv = MeanVoltage, wid=Subject, within=c(GrammarCondition, Electrode))
EEG_ANOVA_resultLL$ANOVA
```

Effect	...	...	F	p	p<.05	ges
<chr>	<dbl×dbl>		<dbl>	<dbl>	<chr>	<dbl>
2 GrammarCondition	1	34	0.4391503	5.120006e-01		0.0006188568
3 Electrode	9	306	3.6757142	2.184057e-04	*	0.0457980035
4 GrammarCondition:Electrode	9	306	4.9914611	2.789725e-06	*	0.0642045237
3 rows						

Hide

```
LL_descriptives <- ANOVA_EEG_dataL %>%
  filter(StimulusCondition %in% 'Language') %>%
  group_by(Electrode, GrammarCondition) %>%
  summarize(Mean = mean(MeanVoltage), SD = sd(MeanVoltage))
LL_descriptives
```

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
33	Grammatical	0.5925397	2.1693547
33	Ungrammatical	-1.3323300	1.6860075
39	Grammatical	0.2447620	2.1452335
39	Ungrammatical	-1.0649484	1.5216531

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
42	Grammatical	0.4474898	2.0987331
42	Ungrammatical	0.1022684	1.5107622
45	Grammatical	-0.1213247	3.2291021
45	Ungrammatical	-0.8190607	1.3187334
70	Grammatical	-0.6599320	2.7207043
70	Ungrammatical	0.4524363	2.1794476

1-10 of 20 rows   Previous **1** 2 Next

ANOVA: Lateral Electrodes, Music

Below is a 2 by 10 factor repeated-measures ANOVA of the mean voltage produced by each participant for music stimuli. EEG data was filtered to the time window of interest, 500 to 800 ms, and the ten lateral electrodes, 33, 39, 42, 45, 70, 83, 93, 108, 115, and 122. Again, no effect of grammar condition was found, which was inconsistent with Patel's results.

Hide

```
EEG_ANOVA_resultLM <- ezANOVA((ANOVA_EEG_dataL %>% filter(StimulusCondition == 'Music'
)), dv = MeanVoltage, wid=Subject, within=c(GrammarCondition, Electrode))
EEG_ANOVA_resultLM$ANOVA
```

Effect	...	...	F	p	p<.05	ges
<chr>	<dbl×dbl>		<dbl>	<dbl>	<chr>	<dbl>
2 GrammarCondition	1	34	1.116281	2.981702e-01		0.001629466
3 Electrode	9	306	11.103754	5.305451e-15	*	0.127124784
4 GrammarCondition:Electrode	9	306	6.469988	1.968589e-08	*	0.071140597

3 rows

Hide

```
LM_descriptives <- ANOVA_EEG_dataL %>%
  filter(StimulusCondition %in% 'Music') %>%
  group_by(Electrode, GrammarCondition) %>%
  summarize(Mean = mean(MeanVoltage), SD = sd(MeanVoltage))
LM_descriptives
```

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
33	Grammatical	0.29643014	1.446836

Electrode	GrammarCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
33	Ungrammatical	-1.04669561	2.495088
39	Grammatical	-0.51972959	1.598382
39	Ungrammatical	-1.11994381	1.801821
42	Grammatical	-0.06532077	1.020731
42	Ungrammatical	1.09063429	1.253943
45	Grammatical	-0.19593580	1.854793
45	Ungrammatical	-0.85663020	1.837532
70	Grammatical	0.21674344	1.610568
70	Ungrammatical	0.78733646	1.612396

1-10 of 20 rows   Previous   **1**   2   Next

Summary While there was not a significant effect of Grammar Condition in the Lateral Electrodes, we did observe an effect in the Midline Electrodes. This weakens the replication, but there is still evidence for an effect of Grammar Condition on grand averaged waveforms, especially because of the interaction observed.
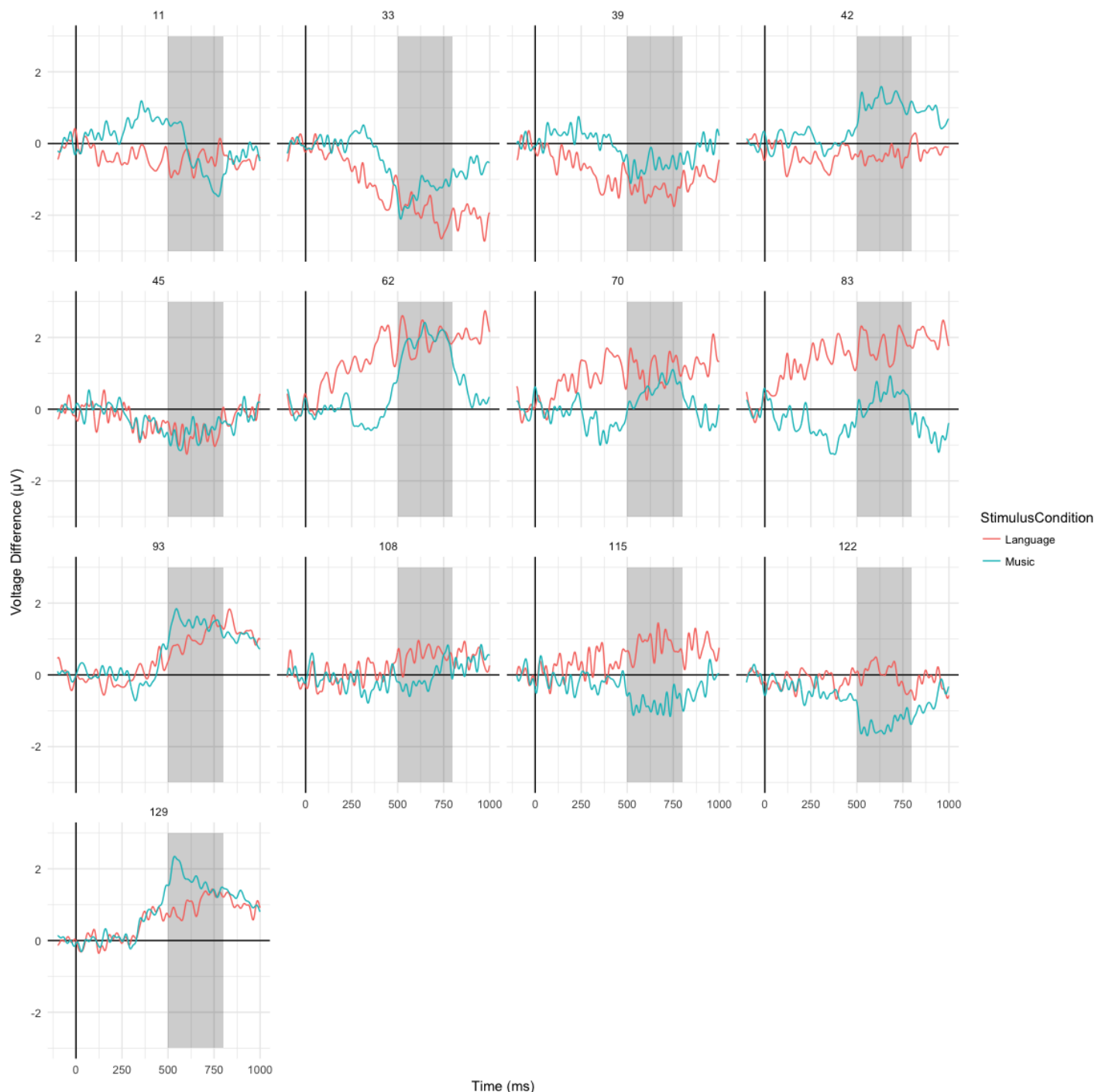
Difference Wave Analysis

Below are direct comparisons of language and music P600 waveforms. The alpha level was set to 0.05. In our OSF registration, this analysis is described as follows:

> Repeated-measures ANOVA, 2 (language v. music) x N (where N = the number of relevant electrodes; see below), of the mean amplitude of the difference wave forms (ungrammatical - grammatical) will be conducted in the 500 to 800 msec separately for lateral and midline electrodes.

The difference waves for the mean voltages between language and music data were calculated (ungrammatical - grammatical), and null data was excluded. A plot of the difference waves is displayed. There was no overall effect or pattern of results in the difference waves, but it does seem that there are elongated positive difference waves for the language condition rather than an isolated spike in the latency window like for the music condition.

[Hide]

```r
difference_waves <- filtered_EEG_data %>%
  group_by(Subject, Electrode, Time, StimulusCondition) %>%
  mutate(DifferenceVoltage = Voltage - lag(Voltage)) %>%
  filter(!is.na(DifferenceVoltage)) %>%
  select(Subject, Time, Electrode, StimulusCondition, DifferenceVoltage) %>%
  ungroup()
grand_average_diff <- difference_waves %>%
  group_by(Electrode, StimulusCondition, Time) %>%
  summarize(GrandDifferenceVoltage = mean(DifferenceVoltage))
ggplot(grand_average_diff, aes(x=Time, y=GrandDifferenceVoltage, color = StimulusConditi
on)) +
  annotate("rect", xmin=500, xmax=800, ymin=-3, ymax=3, fill="black", alpha=0.2) +
  geom_vline(xintercept = 0) +
  geom_hline(yintercept = 0) +
  geom_line() +
  facet_wrap(~Electrode) +
  labs(x='Time (ms)', y='Voltage Difference (µV)') +
  theme_minimal()
```

ANOVA: Midline Electrodes, Language v. Music

The difference wave data was filtered to the midline electrodes (11, 62, and 129) and the designated time window (500 ms to 800 ms). The mean amplitude of the difference waves was defined as the mean of the differences in voltages between language and music. Below is a 2 by 3 factor repeated-measures ANOVA, accompanied by a table summarizing the means and standard deviations of mean difference voltage for each electrode and stimulus condition.

Hide

```
midline_difference_data <- difference_waves %>%
  filter(Time %in% 500:800, Electrode %in% c(11, 62, 129)) %>%
  group_by(Subject, Electrode, StimulusCondition) %>%
  summarize(MeanAmplitude = mean(DifferenceVoltage))
mdd_result <- ezANOVA(midline_difference_data, dv = MeanAmplitude, wid=Subject, within=c
(StimulusCondition, Electrode))
mdd_result$ANOVA
```

Effect	...	...	F	p	p<.05	ges
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<chr>	<dbl>
2 StimulusCondition	1	34	0.2261587	6.374281e-01		0.001426070
3 Electrode	2	68	16.2888392	1.660536e-06	*	0.120421080
4 StimulusCondition:Electrode	2	68	0.3152952	7.306343e-01		0.003329377

3 rows

Hide

```
midline_descriptives <- midline_difference_data %>%
  group_by(Electrode, StimulusCondition) %>%
  summarize(Mean = mean(MeanAmplitude), SD = sd(MeanAmplitude))
midline_descriptives
```

Electrode	StimulusCondition	Mean	SD
<fctr>	<fctr>	<dbl>	<dbl>
11	Language	-0.4721859	2.834194
11	Music	-0.4309371	2.154724
62	Language	1.9811028	4.329985
62	Music	1.9105441	2.752267
129	Language	1.0316035	2.429195
129	Music	1.6861704	1.437693

6 rows

ANOVA: Lateral Electrodes, Language v. Music

The difference wave data was filtered to the lateral electrodes (33, 39, 42, 45, 70, 83, 93, 108, 115, and 122) and the designated time window (500 ms to 800 ms). The mean amplitude of the difference waves was defined as the mean of the differences in voltages between language and music. Below is a 2 by 10 factor repeated-measures ANOVA, accompanied by a table summarizing the means and standard deviations of mean difference voltage for each electrode and stimulus condition. The greatest mean differences were observed in electrodes 42, 115, and 122.

Hide

```
lateral_difference_data <- difference_waves %>%
  filter(Time %in% 500:800, !Electrode %in% c(11, 62, 129)) %>%
  group_by(Subject, Electrode, StimulusCondition) %>%
  summarize(MeanAmplitude = mean(DifferenceVoltage))
ldd_result <- ezANOVA(lateral_difference_data, dv = MeanAmplitude, wid=Subject, within=c
(StimulusCondition, Electrode))
ldd_result$ANOVA
```

Effect <chr>	... <dbl×dbl>	... 	F <dbl>	p <dbl>	p<.05 <chr>	ges <dbl>
2 StimulusCondition	1	34	1.784931	1.904190e-01		0.002027262
3 Electrode	9	306	12.180533	1.816049e-16	*	0.100114976
4 StimulusCondition:Electrode	9	306	2.009149	3.796737e-02	*	0.033709656

3 rows

Hide

```
lateral_descriptives <- lateral_difference_data %>%
  group_by(Electrode, StimulusCondition) %>%
  summarize(Mean = mean(MeanAmplitude), SD = sd(MeanAmplitude))
lateral_descriptives
```

Electrode <fctr>	StimulusCondition <fctr>	Mean <dbl>	SD <dbl>
33	Language	-1.92486965	2.451751
33	Music	-1.34312575	2.290407
39	Language	-1.30971040	2.688846
39	Music	-0.60021422	2.104035
42	Language	-0.34522143	2.590706
42	Music	1.15595506	1.230611
45	Language	-0.69773603	3.581926
45	Music	-0.66069440	2.360228
70	Language	1.11236838	3.887595
70	Music	0.57059302	2.166606

1-10 of 20 rows   Previous   **1**   2   Next

Summary The results of the above two analyses make this study a stronger replication of Patel et al's because it shows that there is no significant effect of the stimulus condition, music or language, on the P600 waveforms, supporting the hypothesis that there may be shared neural resources in processing music and language. An interaction effect of Stimulus Condition on Electrode was found only in the lateral electrodes.

Bayes Factor Analysis

Going off the previous two analyses comparing the effect od language vs. music, Bayes Factor ANOVAs were performed to evaluate support for either Patel et al.'s null or alternative hypotheses. This process is outlined in our OSF registration:

> In the absence of any strong prior beliefs about the size of a difference between music and language P600s, we will use the default prior of the BayesFactor R package (r scaling factor = 0.5) to represent the prior on the fixed effects in a Bayesian repeated-measures ANOVA… We will not include any specific cut-off criterion and instead will report the BF as relative odds in favor of one of the models.

Midline Electrodes, Bayesian Model ANOVA of Language v. Music

Based on the following Bayes Factor ANOVA, the StimulusCondition + Subject model is about 6 times less likely than the Subject model alone, demonstrating that stimulus condition (music vs. language) is not a significant source of variability in the data.

Hide

```
set.seed(10)
midline_difference_data <- as.data.frame((midline_difference_data))
mid_Bayes_ANOVA <- anovaBF(MeanAmplitude ~ StimulusCondition * Electrode + Subject, data
 = midline_difference_data, whichRandom = 'Subject')
```

```
    |

    |
      |    0%
    |

    |=================
      |   20%
    |

    |===================================
      |   40%
    |

    |====================================================
      |   60%
    |

    |=================================================================
      |   80%
    |

    |=======================================================================
====|  100%
```

```
summary(mid_Bayes_ANOVA)
```

```
Bayes factor analysis
--------------
[1] Electrode + Subject                                              : 18094.79
±0.8%
[2] StimulusCondition + Subject                                      : 0.1656667
±0.8%
[3] Electrode + StimulusCondition + Subject                          : 3121.233
±2.09%
[4] Electrode + StimulusCondition + Electrode:StimulusCondition + Subject : 373.0871
±1.96%


Against denominator:
  MeanAmplitude ~ Subject
---
Bayes factor type: BFlinearModel, JZS
```

Lateral Electrodes, Bayesian Model ANOVA of Language v. Music

A similar effect was found in the lateral electrodes:

```
lateral_difference_data <- as.data.frame((lateral_difference_data))
set.seed(10)
lat_Bayes_ANOVA <- anovaBF(MeanAmplitude ~ StimulusCondition * Electrode + Subject, data
 = lateral_difference_data, whichRandom = 'Subject')
```

```
    |

    |
      |    0%
    |


      |==================
      |   20%
    |


      |=================================
      |   40%
    |


      |================================================
      |   60%
    |


      |=============================================================
      |   80%
    |


      |===========================================================================
====|  100%
```

Hide

```
summary(lat_Bayes_ANOVA)
```

```
Bayes factor analysis
--------------
[1] Electrode + Subject                                             : 1806580094
±0.29%
[2] StimulusCondition + Subject                                     : 0.1527808
±0.81%
[3] Electrode + StimulusCondition + Subject                         : 304881734
±2.08%
[4] Electrode + StimulusCondition + Electrode:StimulusCondition + Subject : 1244916765
±1.32%


Against denominator:
  MeanAmplitude ~ Subject
---
Bayes factor type: BFlinearModel, JZS
```

Summary The above two Bayes Factor ANOVA results provide relative support for the null hypothesis because the StimulusCondition + Subject model is significantly less likely (by a factor of about 6) to be responsible for the variability in the data compared to the lone Subject model, indicating that there is little effect of stimulius codition (language or music) on the ERP waveforms.

Discussion

Considering all of the above statistical tests, we conclude that overall this was a successful replication of Patel et al.'s original study. The first six ANOVAs were based on Patel et al.'s analysis methods, but we did not include all three of the time windows that they used in the original study. This is because the ERP component of interest, the P600, should show up in the 500 ms to 800 ms window. Another difference in analysis is our inclusion of Bayesian statistics. Many of our findings, though we had differences in the levels of our grammar conditions, are consistent with Patel. et al.'s, including: main interaction effect between grammar condition and electrode in all four critical conditions (language-midline, music-midline, language-lateral, music-lateral), an effect of grammar condition on the average voltages in the midline electrodes, stronger significant effects (where present) in posterior electrodes, and rationale to fail to reject the null hypothesis in the direct comparison between language and music on P600 waveforms. While the support for the null hypothesis is relative based on the music v. language ANOVAs, we accept them as they are sufficiently powered; our effect size was more than twice that of the original study. The Bayes models served as a more meaningful analysis of the variance in the data because they do not rely on power; this further corroborates the null claim concluded by Patel et al. Still, this is only relatively reliable and does not provide definitive support, because though the Bayes Factor for Stimulus Condition is less likely than the Subject model, the degree to which they differ, if it were greater, would provide even more support.

While there is much support for framing this study as a strong replication of Patel et al., there are some drawbacks to address. An interaction effect between Stimulus Condition and Electrode was found only in the lateral electrodes, which is not consistent with Patel et al.'s findings; it appears that the greater mean differences in certain electrodes (42, 115, 122) may be responsible for this observation. Similarly, no effect of grammar condition was found in the lateral electrodes, but there was in Patel et al.'s study. Additionally, Patel et al. reported to observe peak P600 amplitudes in the language condition in a "lagged" manner, between 800 ms and 900 ms (1998, p. 722), but our peak amplitudes were observed in the 500 ms to 800 ms time window. We also decided to omit the hemispheric analysis that Patel et al. used in some of their statistical tests to observe interaction effects. These differences in results and parameters weaken the replication to an extent, but in the broader sense this study provides support for the central theories that Patel et al. address: that there are structural similarities in spoken language and music, and that structurally incongruous phrases elicit the P600 ERP component.

References

Nosek, B. (2017, December 8). Why are we working so hard to open up science? A personal story. [Blog post]. Retrieved from https://cos.io/blog/why-are-we-working-so-hard-open-science-personal-story/ (https://cos.io/blog/why-are-we-working-so-hard-open-science-personal-story/)

Patel, A.D., Gibson, E., Ratner, J., Besson, M., Holcomb, P.J. (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience, 10*:6, pp. 717–733. doi:10.1162/089892998563121 (doi:10.1162/089892998563121)

Patel, A.D. (2013). Sharing and Nonsharing of Brain Resources for Language and Music. In M.A. Arbib (Ed.), *Language, music, and the brain* (pp. 329-355). Cambridge, MA: MIT Press. doi:978-0-262-01810-4 (doi:978-0-262-01810-4)

Power, A. (2016, November 23). Registered reports: what are they and why are they important? [Blog post]. Retrieved from https://blogs.royalsociety.org/publishing/registered-reports-what-are-they-and-why-are-they-important/ (https://blogs.royalsociety.org/publishing/registered-reports-what-are-they-and-why-are-they-important/)

Srivastava, S. (2014, July 1). Some thoughts on replication and falsifiability: Is this a chance to do better? [Blog post]. Retrieved from https://hardsci.wordpress.com/2014/07/01/some-thoughts-on-replication-and-falsifiability-is-this-a-chance-to-do-better/ (https://hardsci.wordpress.com/2014/07/01/some-thoughts-on-replication-and-falsifiability-is-this-a-chance-to-do-better/)