STAT427 Final Project

Yilun Fu, Yuhang Gong, Chunjiang Li, Zixuan Wang

May 12, 2022

Abstract

This is a report corresponding to the research on the priming effect.

1 Introduction

One of the classic findings in psycholinguistics is structural priming, also known as syntactic priming, or structural repetition. Structural priming is the tendency for speakers to reuse recently experienced structures. Bock (1986a) gave experimental participants pictures that can be described with either of the two kinds of dative structures (double objects, "The woman handed the boy the paint brush," versus prepositional datives, "The woman handed the paint brush to the boy"). Participants described these pictures after saying an unrelated prime sentence that used either a double-object or prepositional dative structure. Priming was seen in the tendency for speakers to use the structure of the prime when describing the picture. Similar effects were seen for other structural alternations such as active transitive sentences ("Lightning is striking the church") versus passives ("The church is struck by lightning"). As such, structural priming provides evidence for a production process that uses structural abstractions during grammatical encoding.

In a word, priming is the psycholinguistic phenomenon whereby exposure to a linguistic structure increases the likelihood of reusing that same structure in subsequent production.

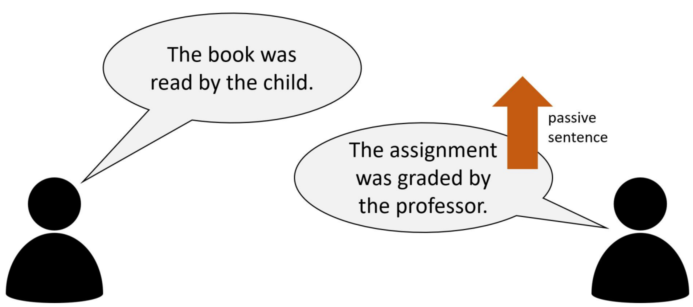

Figure 1: Priming effect example

2 Experiment design

We have three groups and the experiment is designed to examine how priming works in three different groups of Spanish-English bilingual speakers.

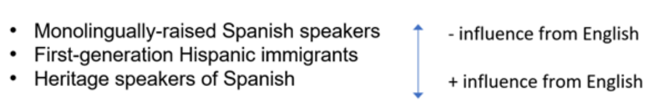

Figure 2: Three groups

Experiment	Construction	Tasks	Language mode
Priming experiment 1	ACC	Pre-test, treatment, post-test	Spanish > Spanish (within)
Priming experiment 2	SPE	Pre-test, treatment, post-test	Spanish > Spanish (within)
Priming experiment 3	SPE	Pre-test, treatment, post-test	English > Spanish (cross)

Figure 3: Research Question 1

The purpose of study question one is to determine if the ACC or SPE architecture has a stronger priming effect.

The second research question aims to find out whether the priming effect is stronger in within-language or cross-linguistic mode.

Experiment	Construction	Tasks	Language mode
Priming experiment 1	ACC	Pre-test, treatment, post-test	Spanish > Spanish (within)
Priming experiment 2	SPE	Pre-test, treatment, post-test	Spanish > Spanish (within)
Priming experiment 3	SPE	Pre-test, treatment, post-test	English > Spanish (cross)

Figure 4: Research Question 2

The third research question is to find out which individual variables are associated with a strong priming effect.

- Bilingual Language Profile (BLP)
- Frequency of language use
- Mean length of utterance (MLU)
- Words per minute
- Vocabulary density (VOCd)

2.1 Data description

Individual and categorical Variables included in experiments:

Individual variable	Details
Bilingual Language Profile	Score from -218 to +218
Frequency of language use	Spanish / English
Mean length of utterance	Spanish / English
Words per minute	Spanish / English
Vocabulary diversity	Spanish / English

Table 1: Individual variables

BLP - The feature BLP indicates the Bilingual Language Profile (BLP). This score ranges from -218 to 218 and it reflects the bilingual dominance of the subject. If the score is negative, the subject is more English dominant and if the score is positive, the subject is more Spanish dominant.

language_use_span - The feature language_use_span indicates the percentage frequency of Spanish usage in the English-Spanish bilingual system. This value ranges from 0 to 100 and the sum of language_use_span and language_use_eng is always 100.

language_use_eng - The feature language_use_eng indicates the percentage frequency of English usage in the English-Spanish bilingual system. This value ranges from 0 to 100 and the sum of language_use_span and language_use_eng is always 100.

Categorical variable	Details
Subjects	124 subjects
Groups	Monolingual / First-gen / Heritage
Phases	Pre-test / Treatment / Post-test
Constructions (in Spanish)	ACC / SPE
Modes	Within / Cross
Target	Yes / No
N items	180 in total / per participant

Table 2: Categorical Variables

MLU_spa - The feature MLU_spa indicates the Mean Length of Utterance (MLU) when using Spanish. The minimum and maximum of this value in this experiment are 7.565 and 18.214, respectively.

MLU_eng - The feature MLU_eng indicates the Mean Length of Utterance (MLU) when using English. The minimum and maximum of this value in this experiment are 5.833 and 16.294, respectively.

Words_Min_spa - The feature Words_Min_spa indicates the Words per Minute when using Spanish. The minimum and maximum of this value in this experiment are 39.389 and 144.793, respectively.

Words_Min_eng - The feature Words_Min_eng indicates the Words per Minute when using English. The minimum and maximum of this value in this experiment are 15.19 and 152.948, respectively.

VOCD_spa - The feature VOCD_spa indicates the Vocabulary Diversity (VOCD) when using Spanish. The minimum and maximum of this value in this experiment are 7.69 and 40.43, respectively.

VOCD_eng - The feature VOCD_eng indicates the Vocabulary Diversity (VOCD) when using English. The minimum and maximum of this value in this experiment are 7.49 and 66.88, respectively.

2.2 Data Cleaning

After we got the data from the client, we made some effort in data cleaning. The problems that we have met and solved in the progress of data cleaning include:

A. The number of subjects in each group is not the same as expected.

Examined the raw experimental data and ensured the number of subjects in each group.

B. The number of tests for several subjects in a certain phase is wrong.

We examined the raw experimental data and found out that several data are incorrectly labeled and corrected the label so that the numbers of each phase are the same as expected.

C. There is one subject that contains two different groups.

We double-checked the group of the subject and made it clear which one of them is corrected.

D. There is one subject that contains two different BLP numbers.

We double-checked the BLP of the subject and made it clear which one of them is corrected.

E. The sum of language_use_eng and language_use_span is not 100 for some subjects.

We rescaled the two values to make them have a sum of 100.

3 Methodology

Three research questions must be addressed for this project: To begin, is the priming effect stronger with the ACC or SPE construction? Second, is the priming effect stronger in cross-linguistic or within-language mode? Third, which individual factors (such as BLP, frequency of language usage, MLU, words per minute, and VOCD) are linked to a strong priming effect? To acquire an overview of the data set, we first performed exploratory data analysis for each of the study questions. Then we created models and used the results to answer the research question. Finally, we did a power analysis to assess the analysis' trustworthiness.

3.1 Methods Used in Exploratory Data Analysis

3.1.1 Chi-Square Test

The Chi-Square test in R is a statistical method used to determine if two categorical variables have a significant correlation between them. *chisq.test(data)* is a function used to perform the test. The two variables are selected from the same population. Furthermore, these variables are then categorized as Male/Female, Red/Green, Yes/No, etc.

Like all statistical tests, we assume this test as a null hypothesis and an alternative hypothesis. For the null hypothesis, we assume the two variables are independent; and for the alternative hypothesis, we assume the two variables relate to each other. We reject the null hypothesis if the p-value that comes out in the result is less than a predetermined significance level, which is 0.05 usually, then we reject the null hypothesis.

$$H_0 : \text{The two variables are independent}$$
$$H_a : \text{The two variables relate to each other}$$

3.1.2 VIF

Multicollinearity occurs when two or more independent variables are highly correlated with one another in a regression model. It will mislead the individual effects of the independent variables on the dependent variable. Therefore, we can use VIF (Variable Inflation Factors) to detect if collinearity exists. VIF score of an independent variable represents how well the variable is explained by other independent variables.

$$VIF = \frac{1}{1 - R^2}$$

It means the closer the R^2 value is to 1, the higher the value of VIF can be and the more likely the multicollinearity will occur with the particular independent variable.

3.1.3 Correlation test

Correlation test is used to evaluate the association between two or more variables. The famous method is Pearson correlation. The formula is

$$r = \frac{\sum(x - \overline{x})(y - \overline{y})}{\sqrt{\sum(x - \overline{x})^2 \sum(y - \overline{y})^2}}$$

If the p-value is less than 0.05, then the correlation between x and y is significant.

3.2 Methods used in Modeling

3.2.1 Generalized linear mixed model fit by maximum likelihood

Generalized linear mixed models (or GLMMs) are an extension of linear mixed models to allow response variables from different distributions, such as binary responses. Alternatively, GLMMs can be applied as an extension of generalized linear models (e.g., logistic regression) to include both fixed and random effects (hence mixed models). And fitting GLMMs via maximum likelihood (as via AIC) involves integrating over the random effects. In fact, those integrals cannot be expressed in analytical form. We can use various approximate methods, one of the famous forms is the Laplace approximation.

3.2.2 Stepwise

Stepwise regression is the step-by-step iterative construction of a regression model that involves the selection of independent variables to be used in a final model. It involves adding or removing potential explanatory variables in succession and testing for statistical significance after each iteration.

The forward selection approach starts with nothing and adds each new variable incrementally, testing for statistical significance.

The backward elimination method begins with a full model loaded with several variables and then removes one variable to test its importance relative to overall results.

3.2.3 Methods used in Power Analysis

Power is defined as the probability of not making a Type II error. Mathematically, power is $1-\beta$. The power of a hypothesis test is between 0 and 1; if the power is close to 1, the hypothesis test is very good at detecting a false null hypothesis. Beta is commonly set at 0.2 but may be set by the researchers to be smaller.

In our project, we use *pwr.f2.test(u, v, f2, sig.level)* to test for the general linear model. The F test has a numerator and denominator degrees of freedom. u, is the number of coefficients you'll have in your model (minus the intercept). v, is the number of error degrees of freedom: $v = n - u - 1$. f2, is the effect size, equals to $\frac{R^2}{1-R^2}$, where R^2 is the coefficient of determination, aka the "proportion of variance explained".

And in Generalized mixed-effect models, we use *r.squaredGLMM()* to calculate pseudo-R-squared.

4 Research Questions

4.1 Research Question one

The purpose of study question one is to determine if the ACC or SPE architecture has a stronger priming effect.

4.1.1 Exploratory Data Analysis (Categorical Features)

First of all, we need to filter the dataset for RQ1 with data only has "mode" equal to "within". After that, let's explore the relationship between "group" and "target" and create a contingency table of them, as well as proportional contingency table:

group	no	yes	Sum
first-gen	2625	255	2880
heritage	3971	829	4800
monolingual	6488	712	7200
Sum	13084	1796	14880

Table 3: Contingency table of "group" and "target"

group	no	yes
first-gen	91.145833	8.854167
heritage	82.729167	17.270833
monolingual	90.111111	9.888889

Table 4: Proportional contingency table of "group" and "target"

For all three categories, the fraction of people who have a priming effect is extremely low, falling below 20 percent of the total. In our study, heritage Spanish speakers had nearly twice the chance of priming effect as monolingually raised Spanish speakers and first-generation Hispanic immigrants.

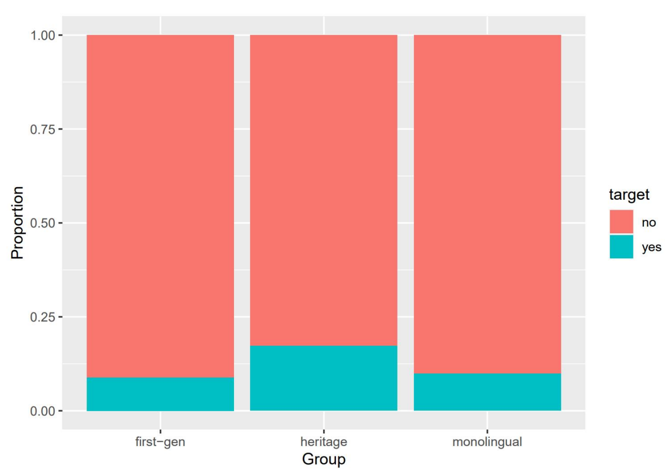

Figure 5: Visualize the proportional contingency table

The Chi-Square test is used to see if there is an association between "Group" and "Target." To put it another way, we may use this tool to see if there is a statistically significant difference in the magnitude of priming effect between the three groups.

```
##
##  Pearson's Chi-squared test
##
## data:  priming_dataset_rq1_sample$group and priming_dataset_rq1_sample$target
## X-squared = 16.993, df = 2, p-value = 0.0002042
```

Figure 6: Chi-Square test

The p value for the Chi-Square test is 0.0002042 which rejects the null hypothesis and concludes there is a statistically significant difference of the priming effect among groups.

Secondly, explore the relationship between "phase" and "target":

phase	no	yes	Sum
post-test	4704	256	4960
pre-test	4779	181	4960
treatment	3601	1359	4960
Sum	13084	1796	14880

Table 5: Contingency table of "phase" and "target"

phase	no	yes
post-test	94.838710	5.161290
pre-test	96.350806	3.649194
treatment	72.600806	27.399194

Table 6: Proportional contingency table of "phase" and "target"

The proportion of having priming effect is very small for both of post-test and pre-test. As for the group of treatment, it has 27.4 percent of having priming effect which is definitely larger than that of post-test and pre-test.

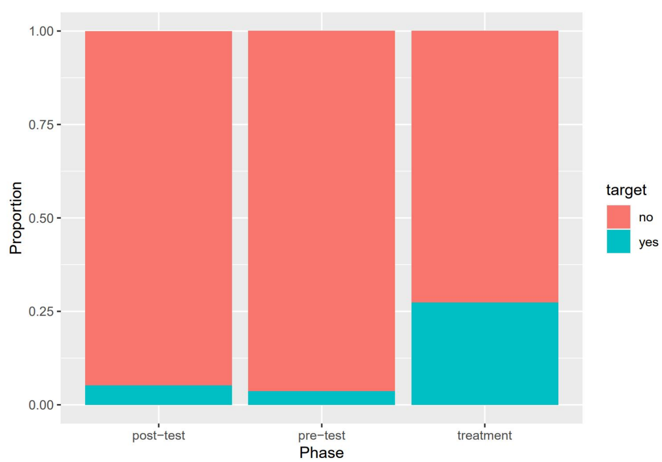

Figure 7: Visualize the proportional contingency table

Again, use the Chi-Square test to test if there is any association between "phase" and "target".

```
##
##  Pearson's Chi-squared test
##

## data:  priming_dataset_rq1_sample$phase and priming_dataset_rq1_sample$target
## X-squared = 116.65, df = 2, p-value < 2.2e-16
```

Figure 8: Chi-Square test

The p value for the Chi-Square test is less than 2.2×10^{-16} which rejects the null hypothesis and concludes there is statistically significant difference of priming effect among different test groups.

Thirdly, explore the relationship between “construction” and “target”. It can be done by the contingency tables as below:

construction	no	yes	Sum
acc	7270	170	7440
spe	5814	1626	7440
Sum	13084	1796	14880

Table 7: Contingency table of ”construction” and ”target”

construction	no	yes
acc	97.715054	2.284946
spe	78.145161	21.854839

Table 8: Proportional contingency table of ”construction” and ”target”

The proportion of having a priming effect is very small for ACC which even looks close to 0. As for the group of SPE, it has 21.9 percent of having a priming effect which is definitely larger than that of ACC.

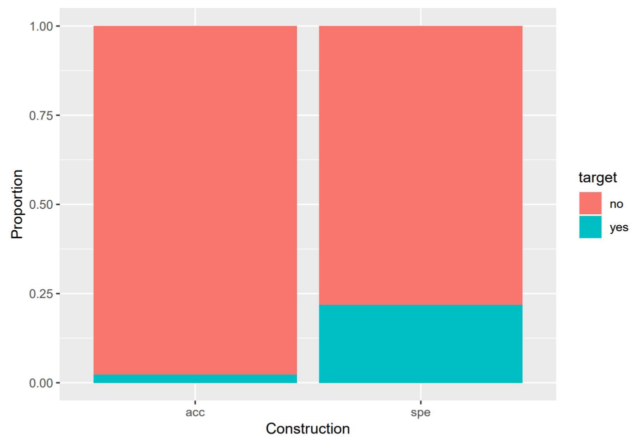

Figure 9: Visualize the proportional contingency table

Again, use the Chi-Square test to test if there is any association between “construction” and “target”.

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  priming_dataset_rq1$construction and priming_dataset_rq1$target
## X-squared = 1340.5, df = 1, p-value < 2.2e-16
```

Figure 10: Chi-Square test

The p value for the Chi-Square test is less than 2.2×10^{-16} which rejects the null hypothesis and concludes there is statistically significant difference of priming effect among different construction groups.

4.1.2 Modeling

Our response variable is a yes/no binary outcome. To fit the data, we utilize a generalized linear mixed model fit by maximum likelihood, which employs the logistic regression method. The estimated variance of random effects is 1.1286 and 0.3499 for subject and n_item respectively. Both of them are greater than 0 so we should keep these random effects.

Intercept, phasetreatment, constructionacc, groupheritage, and the phasepost-test:constructionacc interaction are all significant at the 0.05 level. This indicates that the construction variable should be kept in the model. Constructionacc has a coefficient of -3.7316, indicating that the probability of priming effect for ACC construction should be multiplied by $exp^{-3.7316} = 0.02395448$ when compared to SPE.

```
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  subject (Intercept) 1.1286   1.0624
##  n_item  (Intercept) 0.3499   0.5915
## Number of obs: 14880, groups:  subject, 124; n_item, 120
##
## Fixed effects:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                     -3.7857     0.2847 -13.296  < 2e-16 ***
## phasepost-test                   0.3888     0.2203   1.765   0.0776 .
## phasetreatment                   3.2105     0.2145  14.966  < 2e-16 ***
## constructionacc                 -3.7316     0.4593  -8.125 4.48e-16 ***
## groupheritage                    1.2372     0.2944   4.203 2.64e-05 ***
## groupmonolingual                 0.2704     0.2769   0.977   0.3288
## phasepost-test:constructionacc   1.1400     0.5398   2.112   0.0347 *
## phasetreatment:constructionacc   0.2730     0.5064   0.539   0.5898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

Figure 11: Summary of the model

```
## construction = spe:

##  contrast                 estimate    SE  df z.ratio p.value
##  (pre-test) - (post-test)   -0.389 0.220 Inf  -1.765  0.1815
##  (pre-test) - treatment     -3.211 0.215 Inf -14.966  <.0001
##  (post-test) - treatment    -2.822 0.210 Inf -13.457  <.0001
##
## construction = acc:
##  contrast                 estimate    SE  df z.ratio p.value
##  (pre-test) - (post-test)   -1.529 0.493 Inf  -3.098  0.0055
##  (pre-test) - treatment     -3.484 0.461 Inf  -7.558  <.0001
##  (post-test) - treatment    -1.955 0.294 Inf  -6.648  <.0001
##
## Results are averaged over the levels of: group
## Results are given on the log odds ratio (not the response) scale.
## P value adjustment: tukey method for comparing a family of 3 estimates
```

Figure 12: Marginal means by construction

From the estimate marginal means by construction, we also can find there is a significant difference for (pre-test) - treatment and (post-test) - treatment when construction is spe. And all of the pairwise groups are significant when construction is acc.

```
## phase = pre-test:
##  contrast  estimate    SE  df z.ratio p.value
##  spe - acc     3.73 0.459 Inf   8.125  <.0001
##
## phase = post-test:
##  contrast  estimate    SE  df z.ratio p.value
##  spe - acc     2.59 0.289 Inf   8.954  <.0001
##
## phase = treatment:
##  contrast  estimate    SE  df z.ratio p.value
##  spe - acc     3.46 0.218 Inf  15.876  <.0001
##
## Results are averaged over the levels of: group
## Results are given on the log odds ratio (not the response) scale.
```

Figure 13: Marginal means by phase

No matter what the phase is, there is always has a significant difference between acc and spe.

4.1.3 Power analysis

In order to find out the credibility of our model given current analysis, we do the power analysis for the experiment. The result are shown below:

The result of power analysis indicates that the result we have got for the research question can be credible based on given dataset and statistic model since the power is larger than 0.8.

```
##
##      Multiple regression power calculation
##
##               u = 7
##               v = 116
##              f2 = 0.4913505
##       sig.level = 0.05
##           power = 0.999994
```

Figure 14: Power analysis

4.2 Research Question two

The second research question aims to find out whether the priming effect is stronger in within-language or cross-linguistic mode.

4.2.1 Exploratory Data Analysis (Categorical Features)

The research question 2 focuses on the mode of the subjects and we need to fix the construction to SPE. So before we move on to the next steps of data analysis, we need to filter the dataset to contain only SPE construction.

Let's explore the relationship between "group" and "target" and create a contingency table of them, as well as proportional contingency table:

group	no	yes	Sum
first-gen	2429	451	2880
heritage	3722	1078	4800
monolingual	5989	1211	7200
Sum	12140	2740	14880

Table 9: Contingency table of "group" and "target"

group	no	yes
first-gen	84.34028	15.65972
heritage	77.54167	22.45833
monolingual	83.18056	16.81944

Table 10: Proportional contingency table of "group" and "target"

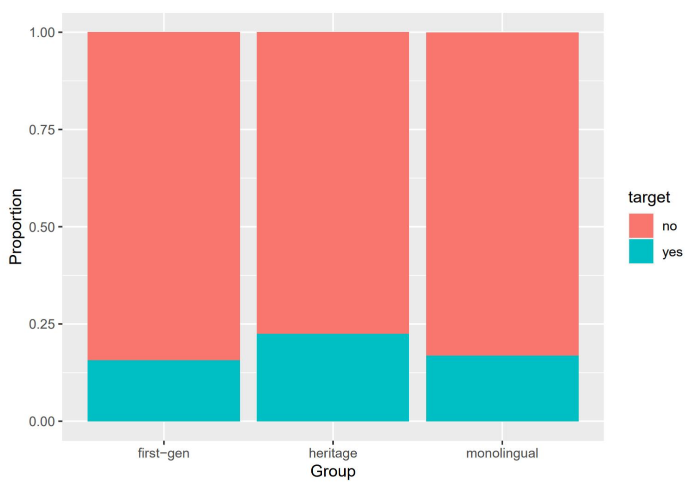

Figure 15: Visualize the proportional contingency table

The proportional contingency table and the visualization indicate that the group of heritage has a significantly higher proportion of priming effect compared to the other two groups.

From the Chi-Square test we can reject the null hypothesis that there is no significant difference between mode and target and come to a conclusion that there is a statistically significant difference between mode and target.

```
## 	Pearson's Chi-squared test
##
## data:  priming_dataset_rq2_sample$group and priming_dataset_rq2_sample$target
## X-squared = 12.131, df = 2, p-value = 0.002321
```

Figure 16: Chi-Square test

Let's explore the relationship between "phase" and "target":

phase	no	yes	Sum
post-test	4531	429	4960
pre-test	4609	351	4960
treatment	3000	1960	4960
Sum	12140	2740	14880

Table 11: Contingency table of "phase" and "target"

phase	no	yes
post-test	91.350806	8.649194
pre-test	92.923387	7.076613
treatment	60.483871	39.516129

Table 12: Proportional contingency table of "phase" and "target"

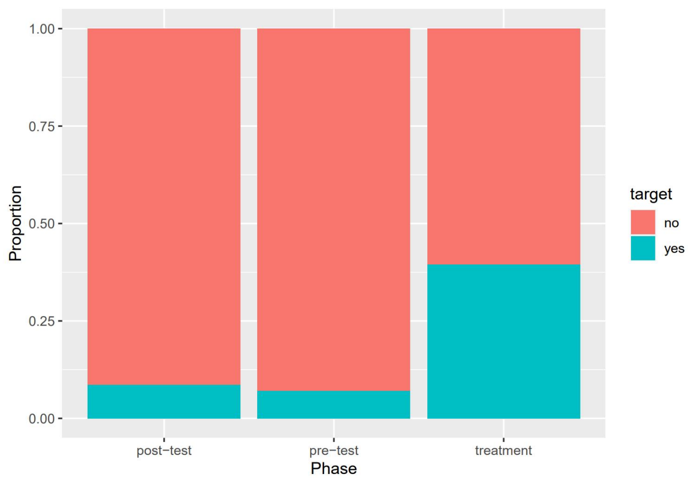

Figure 17: Visualize the proportional contingency table

The proportional contingency table and the visualization indicate that the phase of treatment has a significantly higher proportion of priming effect compared to the other two phases.

```
##
## 	Pearson's Chi-squared test
##
## data:  priming_dataset_rq2_sample$phase and priming_dataset_rq2_sample$target
## X-squared = 151.37, df = 2, p-value < 2.2e-16
```

Figure 18: Chi-Square test

From the Chi-Square test we can reject the null hypothesis that there is no significant difference between phase and target and come to a conclusion that there is a statistically significant difference between "phase" and "target".

Let's explore the relationship between "mode" and "target":

The proportional contingency table and the visualization indicate that the within mode has a significantly higher proportion of priming effect compared to the cross mode. The proportion of cross group to have priming effect is 15 percent, while that of within group is 21.85 percent which is higher.

From the Chi-Square test we can reject the null hypothesis that there is no significant difference between mode and target and come to a conclusion that there is a statistically significant difference between mode and target.

mode	no	yes	Sum
cross	6326	1114	7440
within	5814	1626	7440
Sum	12140	2740	14880

Table 13: Contingency table of "mode" and "target"

mode	no	yes
cross	85.02688	14.97312
within	78.14516	21.85484

Table 14: Proportional contingency table of "mode" and "target"

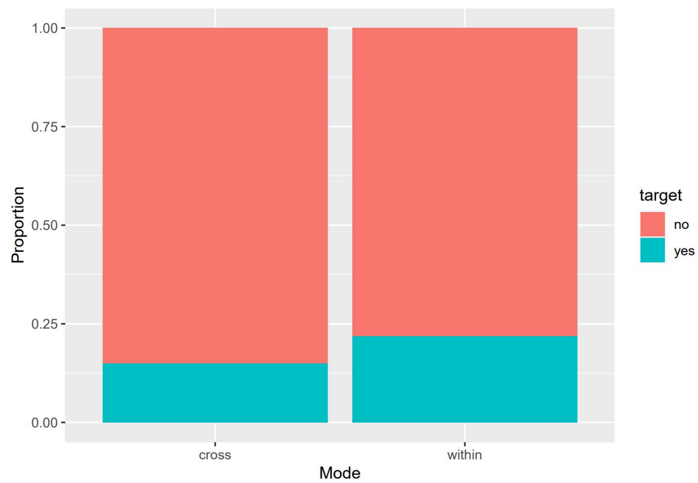

Figure 19: Visualize the proportional contingency table

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  priming_dataset_rq2_sample$mode and priming_dataset_rq2_sample$target
## X-squared = 13.736, df = 1, p-value = 0.0002103
```

Figure 20: Chi-Square test

4.2.2 Modeling

We utilized a GLMM model to fit the dataset to better answer research question 2. "target" is the model's dependent variable, whereas phase, mode, group, subject, and n_item are the independent variables. We also assumed there would be a cross-relationship between phase and group, hence subject and n_item were chosen as random variables. We used R to run the model fitting and got the following results:

```
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  subject (Intercept) 1.212    1.1010
##  n_item  (Intercept) 0.419    0.6473
## Number of obs: 14880, groups:  subject, 124; n_item, 120
##
## Fixed effects:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)               -3.4822     0.2905 -11.988  < 2e-16 ***
## phasepost-test             0.4071     0.2373   1.716 0.086206 .
## phasetreatment             3.2150     0.2301  13.973  < 2e-16 ***
## modecross                  0.0161     0.2410   0.067 0.946754
## groupheritage              0.5716     0.2959   1.932 0.053389 .
## groupmonolingual           0.1071     0.2773   0.386 0.699230
## phasepost-test:modecross  -0.2149     0.3365  -0.639 0.523012
## phasetreatment:modecross  -1.0860     0.3237  -3.355 0.000794 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 21: Summary of the model

The effect of "phase" on "target" is significant and the effect of "group" on "target" is slightly significant. The effect of "mode" is not significant enough but when the "phase" is "treatment", the effect of "mode" is very significant. The coefficient of "phasetreatment:modecross" is negative, which indicates that there is less probability that priming effect will happen when "mode" is cross.

```
## mode = within:
##  contrast                  estimate    SE  df z.ratio p.value
##  (pre-test) - (post-test)   -0.407 0.237 Inf  -1.716  0.1992
##  (pre-test) - treatment     -3.215 0.230 Inf -13.973  <.0001
##  (post-test) - treatment    -2.808 0.225 Inf -12.466  <.0001
##
## mode = cross:
##  contrast                  estimate    SE  df z.ratio p.value
##  (pre-test) - (post-test)   -0.192 0.239 Inf  -0.805  0.6996
##  (pre-test) - treatment     -2.129 0.230 Inf  -9.256  <.0001
##  (post-test) - treatment    -1.937 0.228 Inf  -8.512  <.0001
##
## Results are averaged over the levels of: group
## Results are given on the log odds ratio (not the response) scale.
## P value adjustment: tukey method for comparing a family of 3 estimates
```

Figure 22: Estimate marginal means by mode

There is a significant difference for pre-test - treatment and post-test - treatment when "mode" is within and there is a significant difference for pre-test - treatment and post-test - treatment when "mode" is cross.

```
## phase = pre-test:
##  contrast       estimate    SE  df z.ratio p.value
##  within - cross  -0.0161 0.241 Inf  -0.067  0.9468
##
## phase = post-test:
##  contrast       estimate    SE  df z.ratio p.value
##  within - cross   0.1988 0.235 Inf   0.847  0.3969
##
## phase = treatment:
##  contrast       estimate    SE  df z.ratio p.value
##  within - cross   1.0699 0.216 Inf   4.950  <.0001
##
## Results are averaged over the levels of: group
## Results are given on the log odds ratio (not the response) scale.
```

Figure 23: Estimate marginal means by phase

There is a significant difference for within - cross when "phase" is treatment.

From the above result we can see that the effect of "mode" is not significant standing alone but it turns out to be significant when we fit the "phase" to be treatment. We splited the database into the "phase" and plot the proportional contingency tables by different phases.

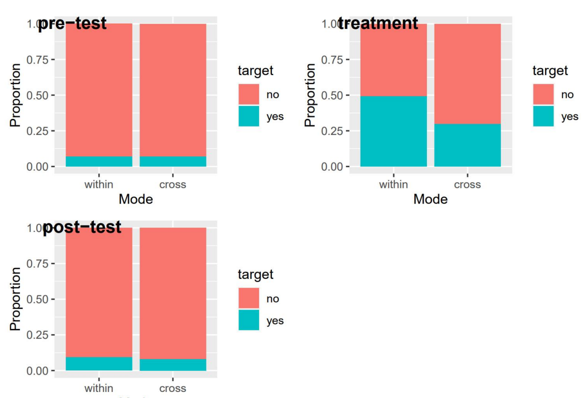

Figure 24: Contingency tables by different phases

We can see that the effect of "mode" on "target" when "phase" is pre-test and post-test are both non-significant. However, the effect of "mode" on "target" when "phase" is treatment is significant. So in this way, the cross effect between "mode" and "phase" should be taken into consideration in our model.

4.2.3 Power analysis

In order to find out the credibility of our model given current analysis, we do the power analysis for the experiment. The result are shown below:

```
##
##      Multiple regression power calculation
##
##               u = 7
##               v = 116
##              f2 = 0.1928426
##       sig.level = 0.05
##           power = 0.9564702
```

Figure 25: Power analysis

The result of power analysis indicates that the result we have got for the research question can be credible based on given dataset and statistic model.

4.3 Research Question three

4.3.1 Exploratory Data Analysis (Numerical Features)

In this section, we focus on these numerical features: "subjects", "n_item", "target", "BLP", "language_use_span", "language_use_eng", "MLU_spa", "Words_Min_spa", "VOCD_spa", "MLU_eng", "Words_Min_eng", "VOCD_eng". First, we use pairplot to check the relationship between these features pairwise.

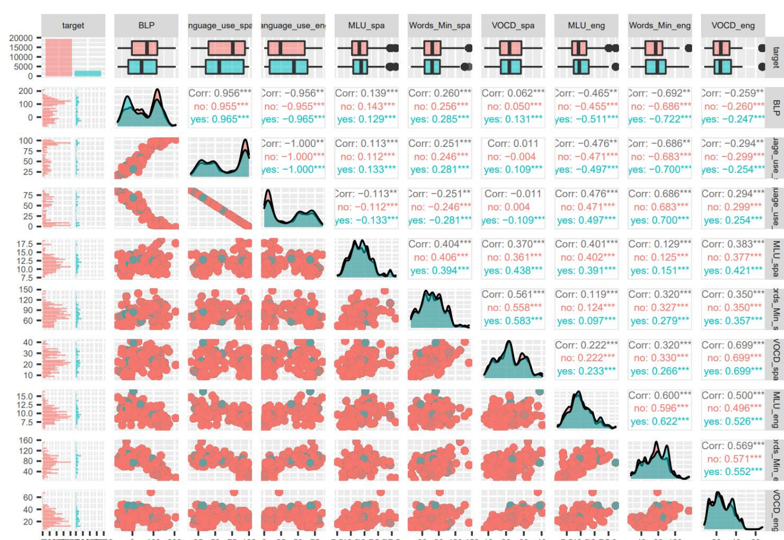

Figure 26: Pairwise plot

Then we check strongly correlated variables and find that there are strong correlations between some variables, and "language_use_span" is exactly correlated with "language_use_eng". Therefore, we decide to remove "language_use_eng".

Moreover, we calculate the mean difference of individual variables between "target" yes and "target" no and get the following form. We found that the difference is small from each other except "BLP_mean".

target	BLP_mean	language_use_span_mean	MLU_spa_mean	Words_Min_spa_mean	VOCD_spa_mean	MLU_eng_mean	Words_Min_eng_mean	VOCD_eng_mean
no	60.20	67.00	11.80	79.00	22.60	9.79	71.09	25.58
yes	48.70	62.60	11.90	77.90	22.30	10.11	74.38	26.31

Figure 27: Mean of variables

4.3.2 Modeling

After exploratory data analysis, we use these data to fit two different models, linear regression (subject level) and generalized linear mixed model fit by maximum likelihood.

To fit linear regression , we need to do the transformation of response variable first as shown in the illustration below:

Original Dataset

subject	n_item	target	BLP	language_use_span	MLU_spa
101	1	no	-7	34	9
101	2	no	-7	34	9
101	3	no	-7	34	9
101	4	no	-7	34	9
101	5	yes	-7	34	9

Transformed Dataset

subject	target_mean	BLP	language_use_span	MLU_spa
101	0.2	-7	34	9

Figure 28: Transformation of response variable

We simplify the dataset to subject level and fit the linear model. We use pairplot to check the relationship between these features pairwise and decide to remove BLP, then fit the model.

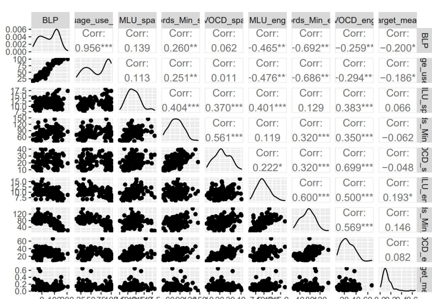

Figure 29: Pair plot

```
##
## Call:
## lm(formula = target_mean ~ BLP + language_use_span + MLU_spa +
##     Words_Min_spa + VOCD_spa + MLU_eng + Words_Min_eng + VOCD_eng,
##     data = priming_dataset_rq3_sbj)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.13954 -0.05191 -0.01441  0.02711  0.50434
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        6.919e-02  8.359e-02   0.828    0.410
## BLP               -3.240e-04  4.669e-04  -0.694    0.489
## language_use_span  2.996e-04  1.092e-03   0.274    0.784
## MLU_spa            3.357e-03  5.308e-03   0.632    0.528
## Words_Min_spa     -8.598e-05  6.767e-04  -0.127    0.899
## VOCD_spa          -1.445e-03  1.902e-03  -0.760    0.449


## MLU_eng            5.101e-03  6.372e-03   0.800    0.425
## Words_Min_eng     -7.185e-05  7.323e-04  -0.098    0.922
## VOCD_eng           5.856e-04  1.428e-03   0.410    0.682
##
## Residual standard error: 0.09572 on 115 degrees of freedom
## Multiple R-squared:  0.0642, Adjusted R-squared:  -0.0009007
## F-statistic: 0.9862 on 8 and 115 DF,  p-value: 0.4505
```

Figure 30: Summary of linear model

According to the p-value, we notice that none of them is significant. So We check the VIF for these variables. "BLP" and "language_use_span" have VIF greater than 10. Therefore, remove one of them ("language_use_span") to address the multicollinearity problem. Then fit the model again.

```
##          BLP language_use_span          MLU_spa     Words_Min_spa
##    13.573668         12.539584         1.661628          2.859228
##     VOCD_spa           MLU_eng     Words_Min_eng          VOCD_eng
##     2.740503          2.137383         5.185029          3.054041
```

Figure 31: VIF

However, none of the predictors is significant. We perform stepwise regression for variable selection to see if we can get a better model. According to the AIC, the model is the best when it only contains "BLP".

This result is same as what we got before.

```
##
## Call:
## lm(formula = target_mean ~ BLP + MLU_spa + Words_Min_spa + VOCD_spa +
##     MLU_eng + Words_Min_eng + VOCD_eng, data = priming_dataset_rq3_sbj)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.14001 -0.04955 -0.01415  0.02734  0.50301
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0842635  0.0627660   1.343    0.182
## BLP           -0.0002156  0.0002480  -0.869    0.386
## MLU_spa        0.0033397  0.0052865   0.632    0.529
## Words_Min_spa -0.0000533  0.0006635  -0.080    0.936
## VOCD_spa      -0.0015131  0.0018783  -0.806    0.422
## MLU_eng        0.0050281  0.0063414   0.793    0.429
## Words_Min_eng -0.0000947  0.0007246  -0.131    0.896
## VOCD_eng       0.0005846  0.0014221   0.411    0.682
##
## Residual standard error: 0.09534 on 116 degrees of freedom
## Multiple R-squared:  0.06359,    Adjusted R-squared:  0.007078
## F-statistic: 1.125 on 7 and 116 DF,  p-value: 0.3521
```

Figure 32: Summary of linear model

```
## Step:  AIC=-584.09
## target_mean ~ BLP
##
##                 Df Sum of Sq    RSS     AIC
## <none>                       1.0807 -584.09
## + MLU_eng        1  0.014392 1.0663 -583.75
## + MLU_spa        1  0.010111 1.0706 -583.25
## + VOCD_spa       1  0.001456 1.0793 -582.26
## + VOCD_eng       1  0.001096 1.0796 -582.21
## + Words_Min_spa  1  0.000132 1.0806 -582.10
## + Words_Min_eng  1  0.000128 1.0806 -582.10
## - BLP            1  0.045232 1.1260 -581.00
```

Figure 33: Stepwise procedure

```
## Call:
## lm(formula = target_mean ~ BLP, data = priming_dataset_rq3_sbj)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.14021 -0.04886 -0.01792  0.03072  0.50147
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.1468994  0.0111761   13.14   <2e-16 ***
## BLP         -0.0002816  0.0001246   -2.26   0.0256 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 34: Summary of stepwise model

Next, we try to fit the Generalized linear mixed model.

First, we standardize the features, then perform the generalized linear mixed model fit by maximum likelihood. From the figure, we find that none of these variables is significant. And after we check VIF scores, "BLP" is 13.875358, "language_use_span" is 12.691955 and others are close to or less than 5. Therefore, we consider that "BLP" and "language_use_span" have multicollinearity and remove "language_use_span", then fit the model again.

```
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)       -3.13628    0.17281 -18.149   <2e-16 ***
## BLP               -0.24429    0.35631  -0.686    0.493
## language_use_span  0.14608    0.34157   0.428    0.669
## MLU_spa            0.05452    0.12357   0.441    0.659
## Words_Min_spa     -0.06522    0.16321  -0.400    0.689
## VOCD_spa          -0.15593    0.15915  -0.980    0.327
## MLU_eng            0.13362    0.13994   0.955    0.340
## Words_Min_eng     -0.05211    0.21849  -0.238    0.811
## VOCD_eng           0.14496    0.16739   0.866    0.386
## ---
```

Figure 35: Summary of GLMM model

```
##               BLP language_use_span           MLU_spa     Words_Min_spa
##         13.875358         12.691955          1.655096          2.860635
##          VOCD_spa           MLU_eng     Words_Min_eng          VOCD_eng
##          2.753717          2.139191          5.202090          3.065345
```

Figure 36: VIF

However, after removing language_use_span and refit the GLMM model we get the same result that none of the predictors is significant at 0.05 level. As for stepwise variable selection, it only contains "BLP" at the end.

4.3.3 Power analysis

For Generalized Mixed-Effect model without the variable "language use span", we calculate Pseudo-R-squared and the power, which is 0.07588516. The power is too small, we get the conclusion that none of these individual variables are associated with a strong priming effect.

5 Summary of result

- Question 1: Is the priming effect stronger with the ACC or with the SPE construction?

The priming effects are less likely to happen when using ACC construction compared to SPE construction.

- Question 2: Is the priming effect stronger in within-language mode or in cross-linguistic mode?

The modecross itself is not significant but the interaction of phasetreatment and modecross is significant. Only in the treatment phase of the experiment, priming effects are less likely to happen in cross mode compared to within mode.

- Question 3: Which individual variables are associated with a strong priming effect (including BLP, frequency of language use, MLU, words per minute and VOCD)?

According to the p-value, none of these is significant and the power also shows that none of these individual variables is associated with a strong priming effect.