Coding Categorical Predictors

for LSA 2013, LI539
Mixed Effect Models

Maureen Gillespie

Department of Psychology
University of New Hampshire

July 5, 2013

Auxiliary lecture 1: Coding of categorical predictors

References

Cohen, & Cohen, (1983). *Applied multiple regression/correlation analysis for the behavioral sciences.*

- Especially chapters 8 & 9

Kaufman, D. & Sweet, R. (1974). Contrast coding in least squares regression analysis. *American Educational Research Journal*, *11*, 359–377.

Serlin, R. C., & Levin, J. R. (1985). Teaching how to derive directly interpretable coding schemes for multiple regression analysis. *Journal of Educational Statistics*, *10*, 223–238.

Wendorf, C. A. (2004). Primer on multiple regression coding: Common forms and the additional case of repeated contrasts. *Understanding Statistics*, *3*, 47–57

How do we treat categorical variables in regression?

- As sets of IVs (code variables)
  - Together they represent the full information from original categories.
- Multiple ways to set up code variables
  - Different ways test different predictions
  - These are essentially planned comparisons

How do we treat categorical variables in regression?

- As sets of IVs (code variables)
  - Together they represent the full information from original categories.
- Multiple ways to set up code variables
  - Different ways test different predictions
  - These are essentially planned comparisons

How many coding variables are necessary?

For any grouped/non-continuous IV (**G**) with some number of levels (g), g - 1 coding variables are needed to represent **G**.

- 4 levels $\rightarrow$ 3 coding variables (C_1, C_2, C_3)
- 3 levels $\rightarrow$ 2 coding variables (C_1, C_2)
- 2 levels $\rightarrow$ 1 coding variables (C_1)

NB:

g -1 = # of degrees of freedom (df) of **G**

How do we represent the coding variables?

Common coding systems

- Treatment/Dummy Coding
- Effects/Sum Coding
- Helmert Coding
- Polynomial Coding

NB:

The choice of your coding scheme affects the interpretation of the results for each **individual coding variable***; however, it does not change the* **overall effect** *of the set of coding variables (i.e., model fit and related statistics will not be affected).*

Example Data Set

Lexical Decision Task

- Word status
  - /smok/ = word
  - /plok/ = phonologically legal nonword
  - /lbok/ = phonologically illegal nonword
- Task: Press button if the item is or sounds like an English word.
- DV: RT of response.

Read in Data


```
library(languageR)
library(lme4)
d<-read.table("data/fakedata.txt", header=TRUE)
#renames factors
d$IV1<-ifelse(d$IV1==1, "silent", "noise")
d$IV2<-ifelse(d$IV2==2, "word", ifelse(d$IV2==3, "legal", "illegal"))
d$NoiseCond<-as.factor(d$IV1)
d$WordCond<-as.factor(d$IV2)
d$WordCond<-as.factor(d$WordCond)
d$NoiseCond<-as.factor(d$NoiseCond)
d$Freq<-as.numeric(d$Freq)
head(d)

##   Subject Item    IV1  IV2 Response  RT Freq NoiseCond WordCond
## 1       1    1 silent word        1 762 0.24    silent     word
## 2       1    2 silent word        1 608 5.22    silent     word
## 3       1    3 silent word        1 744 5.10    silent     word
## 4       1    4 silent word        1 609 5.77    silent     word
## 5       1    9 silent word        1 965 2.37    silent     word
## 6       1   10 silent word        1 817 5.80    silent     word
```

Condition Means


```
attach(d)
#RT means for single WordCond variable

RTmeans<-aggregate(RT, list(WordCond), FUN=mean)
#RT means for two variables (WordCond and NoiseCond - used in interactions.)
RTmeansWN<-aggregate(RT, list(WordCond, NoiseCond), FUN = mean)
RTmeans

##   Group.1      x
## 1 illegal 1315.5
## 2   legal  969.7
## 3    word  731.7

RTmeansWN

##   Group.1 Group.2      x
## 1 illegal   noise 1462.3
## 2   legal   noise 1035.5
## 3    word   noise  738.3
## 4 illegal  silent 1168.7
## 5   legal  silent  903.9
## 6    word  silent  725.1
```

Example Data Set

Lexical Decision Task

- Does word status affect the time to make responses?

We'll run linear mixed-effect models testing this general question with different coding schemes.

- *One fixed effect (WordCond) and two random effects (Subject and Item intercepts)*

Treatment Coding

Compares other groups to a reference group.

- Considerations for choosing a reference group
    - Useful comparison (e.g., control, predicted highest or lowest)
    - Well-defined group (e.g., not a catch-all category)
    - Should **not** have small *n* compared to other groups
- Intercept represents the reference group mean.

Treatment Coding

Imagine the question we're interested in is whether responses to each of the nonword conditions differ from the *word* condition.

Question

What level should we choose as a reference group?

Treatment Coding

Imagine the question we're interested in is whether responses to each of the nonword conditions differ from the *word* condition.

We should choose *word* as our reference group.

- Reference group receives a value of 0 for all coding variables (C_i)
- Each other level receives 1 in one of the coding variables

Levels	$\mathbf{C}_1$	$\mathbf{C}_2$
word	0	0
legal	1	0
illegal	0	1

$\mathbf{C}_1$ *tests legal against word*
$\mathbf{C}_2$ *tests illegal against word*

Treatment Coding


```
#Create column for treatment coded WordCond
d$WordCond.Treatment<-d$WordCond


# R automatically assigns levels alphabetically, this isn't always
#what you'll want, so you can reassign the order of the levels as
#shown below...


d$WordCond.Treatment<-factor(d$WordCond.Treatment, levels=c("word","legal","illegal"))

#This line reorders levels to put "word" in baseline position (1st in list)

# R's default is to set coding scheme to Treatment, so here you
#don't need to do anything else now that the levels are ordered
#appropriately.

#More generally, if you just want to specify which level is the
#baseline you can do the following:

 #contrasts(d$WordCond.Treatment)<-contr.treatment(3, base=3)

#This says set the contrasts to treatment coding with 3 levels,
#with the 3rd level being the base condition

lin.Treatment<-lmer(RT ~ WordCond.Treatment + (1|Subject) + (1|Item), data=d) #linear model
```

Problem 1

Now, let's imagine that you wanted to see RTs for phonologically legal items (English words and legal nonwords) differ from the RTs for phonologically illegal items.

Task

- Choose a base group that tests this question.
- Set up this coding scheme in R.
- Run the model and interpret the coefficients.

Treatment Coding

```
#Create column for treatment coded WordCond
d$WordCond.Treatment<-d$WordCond



d$WordCond.Treatment<-factor(d$WordCond.Treatment, levels=c("word","legal","illegal"))


 contrasts(d$WordCond.Treatment)<-contr.treatment(3, base=3)


Problem1<-lmer(RT ~ WordCond.Treatment + (1|Subject) + (1|Item), data=d) #linear model
```

Treatment Coding


```
Problem1

## Linear mixed model fit by REML
## Formula: RT ~ WordCond.Treatment + (1 | Subject) + (1 | Item)
##    Data: d
##    AIC   BIC logLik deviance REMLdev
##  15582 15612  -7785    15590   15570
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.39e-06 1.18e-03
##  Item     (Intercept) 3.03e+01 5.50e+00
##  Residual             4.42e+04 2.10e+02
## Number of obs: 1152, groups: Subject, 24; Item, 8
##
## Fixed effects:
##                     Estimate Std. Error t value
## (Intercept)           1315.5       10.9   120.7
## WordCond.Treatment1   -583.8       15.2   -38.5
## WordCond.Treatment2   -345.8       15.2   -22.8
##
## Correlation of Fixed Effects:
##             (Intr) WrC.T1
## WrdCnd.Trt1 -0.696
## WrdCnd.Trt2 -0.696  0.500
```

- **Intercept**: Illegal nonword mean RT is 1316ms.
- **C_1**: Legal nonwords are responded to 346ms faster than illegal nonwords.
- **C_2**: Words are responded to 584ms faster than illegal nonwords.

Effects Coding

Compares mean of a single group to the grand mean.

- Usually useful for unordered experimental groups
- *Base group* is chosen
  - Choose "least" interesting group
- Sum of the contrast weights of the coding variables always equals 0.
- Intercept represents the grand mean.

Effects Coding

Imagine that we choose *word* as our base group.

- Base group receives a value of -1 for all coding variables (C_i)
- Each other level receives 1 in one of the coding variables

Levels	$\mathbf{C}_1$	$\mathbf{C}_2$
word	-1	-1
legal	0	1
illegal	1	0

$\mathbf{C}_1$ *is the difference between illegal and grand mean.*
$\mathbf{C}_2$ *is the difference between the legal and grand mean.*

Effects Coding


```
d$WordCond.Effects<-d$WordCond

contrasts(d$WordCond.Effects)<-cbind("illegal.v.GM"= c(1, 0, -1),  "legal.v.GM"= c(0, 1, -1))

#renames Cis to give indication of what is being tested...
#C1 = illegal vs. grandmean, C2= legal vs.grandmean

lin.Effects<-lmer(RT ~ WordCond.Effects + (1|Subject) + (1|Item), data=d)
```

Effects Coding


```
## Linear mixed model fit by REML
## Formula: RT ~ WordCond.Effects + (1 | Subject) + (1 | Item)
##    Data: d
##    AIC   BIC logLik deviance REMLdev
##  15584 15614  -7786    15590   15572
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.39e-06 1.18e-03
##  Item     (Intercept) 3.03e+01 5.50e+00
##  Residual             4.42e+04 2.10e+02
## Number of obs: 1152, groups: Subject, 24; Item, 8
##
## Fixed effects:
##                             Estimate Std. Error t value
## (Intercept)                  1005.65       6.49   154.9
## WordCond.Effectsillegal.v.GM   309.87       8.76    35.4
## WordCond.Effectslegal.v.GM     -35.95       8.76    -4.1
##
## Correlation of Fixed Effects:
##                     (Intr) WrdCnd.Effctsll..GM
## WrdCnd.Effctsll..GM  0.000
## WrdCnd.Effctslg..GM  0.000 -0.500
```

- **Intercept**: Grand mean RT is 1006ms.
- $\mathbf{C}_1$: Illegal nonwords are responded to $\sim$ 310ms slower than the grand mean.
- $\mathbf{C}_2$: Legal nonwords are responded to $\sim$ 37ms faster than the grand mean.

Orthogonal Contrast Coding

Goal of these coding systems is to allow each coding variable (C_i) to

- capture unique portions of the variance (i.e., orthogonal).
- test specific, theory-guided hypotheses (i.e., planned comparisons).

Constructing Orthogonal Contrast Codes (Cohen & Cohen, 1984)

- **Rule 1.** The sum of the weights across each code variable (C_i) must equal 0.
- **Rule 2.** The sum of the products of each pair of code variable (C_1, C_2) must equal 0.
  - When group sizes are equal, this ensures that contrast codes are orthogonal (i.e., do not capture overlapping portions of the variance).
- **Rule/Suggestion 3.** The difference between the value of the set of positive weights and the value of the set of negatives weights should equal 1.
  - Allows each unstandardized β to correspond to the difference between the unweighted means of the groups involved in the contrast.

Treatment and Effects Coding variables are NOT orthogonal!

Treatment Coding

Levels	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{C}_1\mathbf{C}_2$
word	0	0	$= 0$
legal	1	0	$= 0$
illegal	0	1	$= 0$
sum	**1**	**1**	0

Violates Rule 1

Effects Coding

Levels	$\mathbf{C}_1$	$\mathbf{C}_2$	$\mathbf{C}_1\mathbf{C}_2$
word	-1	-1	$= 1$
legal	1	0	$= 0$
illegal	0	1	$= 0$
sum	0	0	**1**

Violates Rule 2

Whenever possible and predictions allow it, use orthogonal coding schemes.

Helmert Coding

Tests one level of a factor against all previous levels.

- Useful for ordinal variables
- Example comparisons
  - Does Level 1 differ from Level 2?
  - Does Level 1 differ from the mean of Levels 2 & 3?
- Intercept represents the grand mean.

Helmert Coding (Regression-style)

Are real English words more quickly or more often perceived as words than nonwords?

Are listeners sensitive to phonotactics of nonwords such that they more quickly and more often perceive phonologically legal nonwords as words than phonologically illegal nonwords?

Levels	**C_1**	**C_2**
word	0	2/3
legal	1/2	-1/3
illegal	-1/2	-1/3

C_2 *tests legal against illegal)*
C_1 *tests word against mean of legal and illegal (i.e., word vs. nonword)*

NB:

R does not automatically assign weights that satisfy Rule/Suggestion 3.

Helmert Coding


```
d$WordCond.Helm.Reg<-d$WordCond

contrasts(d$WordCond.Helm.Reg)<-
  cbind("leg.vs.ill"= c(-.5, .5, 0),
        "word.vs.nons"=c (-(1/3), -(1/3), (2/3))
  )

#renames Cis to give indication of what is being tested... C1 = illegal vs. legal,
#C2= word vs.nonwords(mean of other two levels)

lin.Helm.Reg<-lmer(RT ~ WordCond.Helm.Reg + (1|Subject) + (1|Item), data=d)
```

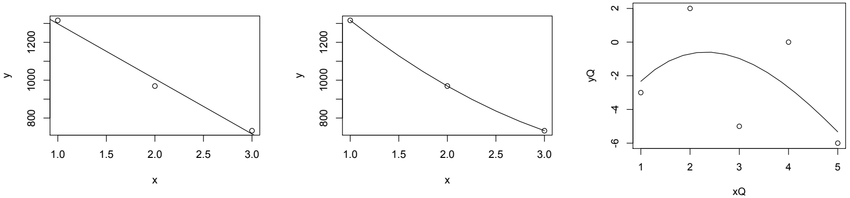
Helmert Coding


```
## Linear mixed model fit by REML
## Formula: RT ~ WordCond.Helm.Reg + (1 | Subject) + (1 | Item)
##    Data: d
##    AIC   BIC logLik deviance REMLdev
##  15582 15612  -7785    15590   15570
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.39e-06 1.18e-03
##  Item     (Intercept) 3.03e+01 5.50e+00
##  Residual             4.42e+04 2.10e+02
## Number of obs: 1152, groups: Subject, 24; Item, 8
##
## Fixed effects:
##                              Estimate Std. Error t value
## (Intercept)                   1005.65       6.49   154.9
## WordCond.Helm.Regleg.vs.ill   -345.83      15.17   -22.8
## WordCond.Helm.Regword.vs.nons -410.87      13.14   -31.3
##
## Correlation of Fixed Effects:
##                   (Intr) WrdCnd.Hlm.Rgl..
## WrdCnd.Hlm.Rgl.. 0.000
## WrdCnd.Hlm.Rgw.. 0.000  0.000
```

C$_1$: *Phonologically legal nonwords are responded to 346ms faster than phonologically illegal nonwords.*
C$_2$: *English words are responded to 411ms faster than nonwords.*

Polynomial Coding

What if we care about the shape of the effect over a range of ordered levels of our independent variable, rather than differences between group means?



Polynomial Coding

How do we model trends in ordered, *equally-spaced*, categorical variables?

- Linear trend?
- Quadratic trend?
- Higher-level trends?

Can test for g - 1 higher-order trends.

- 2-level factor: Linear (X^1)
- 3-level factor: Linear, Quadratic (X^2)
- 4-level factor: Linear, Quadratic, Cubic (X^3)

NB:

Orthogonal polynomial contrasts can be automatically generated by R for any number of levels using the function contr.poly(n), where n = number of levels of your factor.

Polynomial Coding

How do we model trends in ordered, categorical variables?

```
d$WordCond.Poly <- d$WordCond

contrasts(d$WordCond.Poly)<- contr.poly(3)

lin.Poly<- lmer(RT~WordCond.Poly + (1|Subject) + (1|Item), data=d)
```

C_1 (.L) tests if there is a linear component.
C_2 (.Q) tests if there is a quadratic trend.

Polynomial Coding

```
lin.Poly

## Linear mixed model fit by REML
## Formula: RT ~ WordCond.Poly + (1 | Subject) + (1 | Item)
##    Data: d
##    AIC   BIC logLik deviance REMLdev
##  15583 15613  -7785    15590   15571
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.39e-06 1.18e-03
##  Item     (Intercept) 3.03e+01 5.50e+00
##  Residual             4.42e+04 2.10e+02
## Number of obs: 1152, groups: Subject, 24; Item, 8
##
## Fixed effects:
##                 Estimate Std. Error t value
## (Intercept)      1005.65       6.49   154.9
## WordCond.Poly.L  -412.80      10.73   -38.5
## WordCond.Poly.Q    44.04      10.73     4.1
##
## Correlation of Fixed Effects:
##             (Intr) WC.P.L
## WrdCnd.Pl.L 0.000
## WrdCnd.Pl.Q 0.000  0.000
```

C$_1$: *Significant linear trend.*
C$_2$: *Significant quadratic trend.*

Coding Categorical Predictors

for LSA 2013, LI539
Mixed Effect Models

Maureen Gillespie

Department of Psychology
University of New Hampshire

July 5, 2013

Auxiliary lecture 2: Interactions

Interactions with Categorical Variables

How do we interpret interactions when using categorical variables?

- "Difference of differences" when looking at interactions with two (or more) categorical variables.
- Differences among slopes when looking at interactions with a categorical variable and a continuous variable.
- Choice of coding scheme affects interpretation of βs and intercept

Example Data Set - Categorical × Categorical

Lexical Decision Task

- WordCond: /smok/ > /plok/ > /lbok/
  - /smok/ = word
  - /plok/ = phonologically legal nonword
  - /lbok/ = phonologically illegal nonword
- NoiseCond: Noise vs. Silence

- Task: Press button if the item sounds like or is an English word.
- DV: Response Time (RT)

Mean RT (ms) by Word Condition and Noise Condition

	Word	Legal	Illegal
Noise	738	1035	1462
Silence	725	904	1169

Treatment Coding

Reference group receives a value of 0 for all coding variables (C_i)

- We'll choose *word* as our reference group for WordCond and *noise* as our reference group for NoiseCond

Each other level receives 1 in one of the coding variables

Levels	WC_1	WC_2
word	0	0
legal	1	0
illegal	0	1

Levels	NC_1
noise	0
silent	1

NB:

Intercept represents the mean of the noise-word condition

Treatment Coding Interaction


```
d$NoiseCond.Treatment <- d$NoiseCond

contrasts(d$NoiseCond.Treatment)<-contr.treatment(2)
contrasts(d$WordCond.Treatment)<-contr.treatment(3)


# R's default is to treatment code
l.Treatment.int<-lmer(RT ~
    NoiseCond.Treatment*WordCond.Treatment +
    (1|Subject) + (1|Item), data=d)
```

Treatment Coding Interaction


```
## Linear mixed model fit by REML
## Formula: RT ~ NoiseCond.Treatment * WordCond.Treatment + (1 | Subject) +      (1 | Item)
##    Data: d
##    AIC   BIC logLik deviance REMLdev
##  15314 15359  -7648    15338   15296
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept)     0.0    0.00
##  Item     (Intercept)    90.2    9.49
##  Residual             35588.6  188.65
## Number of obs: 1152, groups: Subject, 24; Item, 8
##
## Fixed effects:
##                                          Estimate Std. Error t value
## (Intercept)                                 738.3       14.0    52.7
## NoiseCond.Treatment2                        -13.2       19.3    -0.7
## WordCond.Treatment2                         297.1       19.3    15.4
## WordCond.Treatment3                         724.0       19.3    37.6
## NoiseCond.Treatment2:WordCond.Treatment2   -118.4       27.2    -4.3
## NoiseCond.Treatment2:WordCond.Treatment3   -280.4       27.2   -10.3
##
## Correlation of Fixed Effects:
##             (Intr) NsC.T2 WrC.T2 WrC.T3 NC.T2:WC.T2
## NsCnd.Trtm2 -0.687
## WrdCnd.Trt2 -0.687  0.500
## WrdCnd.Trt3 -0.687  0.500  0.500
## NC.T2:WC.T2  0.485 -0.707 -0.707 -0.354
## NC.T2:WC.T3  0.485 -0.707 -0.354 -0.707  0.500
```

Interpretation of Treatment Coding Model

	Word	Legal	Illegal
Noise	738	1035	1462
Silence	725	904	1169

- Intercept (738)= Noise-Word Condition
- NoiseCondsilent (-13) = Silence-Word − Noise-Word
- WordCondlegal (297)= Noise-Legal − Noise-Word
- WordCondlegal (724)= Noise-Illegal − Noise-Word
- NCsil*WCleg (-118)= (Silence-Legal − Noise-Legal) − (Silence-Word − Noise-Word)
- NCsil:WCill (-280)= (Silence-Illegal − Noise-Illegal) − (Silence-Word − Noise-Word)

Comparison of models with different coding schemes

Model fit (i.e., `predict(model)`) is identical for all coding schemes

Each complete set of coding variables captures the same overall proportion of the variance in the DV, but the interpretation of each individual β *is different under different coding schemes.*

This means that significance of each individual coefficient can vary depending on the chosen coding scheme; however, overall significance of an effect (equivalent to main effects and interactions in an ANOVA) remains the same.

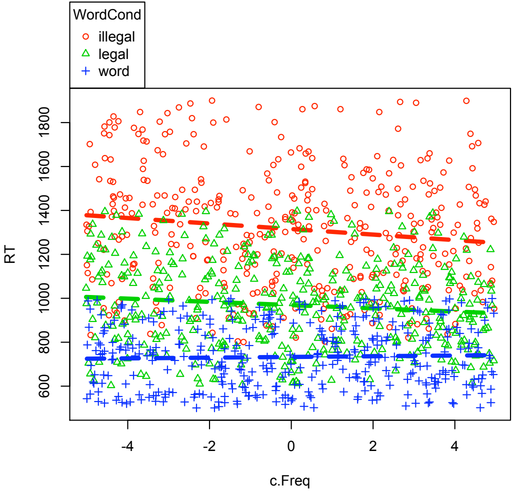
Example Data Set - Continuous $\times$ Categorical

Lexical Decision Task

- c.Freq: Centered log Freq
- WordCond: Word vs. Legal Nonword vs. Illegal Nonword

- Task: Press button if the item sounds like an English word.
- DV: Response Time (RT)

Example Data Set - Continuous × Categorical



Illegal: -12.6ms/unitFreq

Legal: -7.4ms/unitFreq

Word: 1.6ms/unitFreq

Treatment Coding Interaction

Imagine we want to know whether the effect of phoneme frequency differed between words and nonwords.

- Reference group receives a value of 0 for all coding variables (C_i)
  - We'll choose *word* as our reference group for WordCond
- Each other level receives 1 in one of the coding variables

Levels	WC_1	WC_2
word	0	0
legal	1	0
illegal	0	1

Continuous Variable

c.Freq (centered log frequency)

NB:

Intercept represents the mean of the word condition at the mean frequency.

Treatment Coding Interaction


```
d$c.Freq<-d$Freq-mean(d$Freq)

l.Treatment.cont.c<-lmer(RT ~ c.Freq*WordCond.Treatment + (1|Subject) + (1|Item), data=d)
```

Treatment Coding Interaction


```
l.Treatment.cont.c

## Linear mixed model fit by REML
## Formula: RT ~ c.Freq * WordCond.Treatment + (1 | Subject) + (1 | Item)
##    Data: d
##    AIC   BIC logLik deviance REMLdev
##  15559 15604  -7770    15574   15541
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 5.13e-07 7.16e-04
##  Item     (Intercept) 3.63e+01 6.02e+00
##  Residual             4.37e+04 2.09e+02
## Number of obs: 1152, groups: Subject, 24; Item, 8
##
## Fixed effects:
##                            Estimate Std. Error t value
## (Intercept)                  731.73      10.88    67.3
## c.Freq                         1.71       3.69     0.5
## WordCond.Treatment2          238.11      15.09    15.8
## WordCond.Treatment3          583.50      15.09    38.7
## c.Freq:WordCond.Treatment2    -8.92       5.20    -1.7
## c.Freq:WordCond.Treatment3   -14.28       5.23    -2.7
##
## Correlation of Fixed Effects:
##             (Intr) c.Freq WrC.T2 WrC.T3 c.F:WC.T2
## c.Freq      -0.001
## WrdCnd.Trt2 -0.694  0.001
## WrdCnd.Trt3 -0.694  0.001  0.500
## c.Frq:WC.T2  0.001 -0.709 -0.004 -0.001
## c.Frq:WC.T3  0.001 -0.705 -0.001  0.003  0.500
```

Interpretation of Treatment Coding Model

c.Freq: In word condition, for every unit increase in log frequency, RT is 1.7ms slower.

WC2: Legal nonwords are responded to 238ms slower than words.

WC3: Illegal nonwords are responded to 584ms slower than words.

c.Freq*WC2: The frequency effect in the legal condition is non-significantly in the opposite direction of the frequency effect in word condition (legal freq effect: $1.71 + (-8.92) = -7.2$ms/unit Freq).

c.Freq*WC3: The frequency effect in the illegal condition is significantly in the opposite direction of the frequency effect in word condition (illegal freq effect: $1.71 + (-14.28) = -12.6$ms/unit Freq).

But do I have an omnibus interaction?

Usually theoretical predictions do not concern the presence of an omnibus interaction involving multi-level categorical predictors.

- If you're using ANOVA, the presence of an omnibus interaction provides justification to test comparisons of theoretical interest.
- In regression, you can often code your variables in a way to test these specific hypotheses.

If you want to determine whether the omnibus interaction significantly improves model fit, you can use model comparison.

```
m1 <- lmer(NoiseCond*WordCond + (1|Subject) + (1|Item))
m2 <- lmer(NoiseCond+WordCond + (1|Subject) + (1|Item))
anova(m1, m2)
```

Summary and Tips

How you code your predictors determines which hypotheses you are testing.

- Think about what paired comparisons you are interested in testing prior to setting up a coding scheme.
- Choose your baselines carefully, and know what each coding variable is testing.
  - This will help you know which means/differences are being tested and what the sign of your coefficient represents.
- Would a variable you would treat as unordered and categorical in ANOVA be better served as an ordered variable (Helmert, Polynomial), or is it actually continuous?

THANK YOU!!!