A case for systematic sound symbolism in pragmatics: Supporting information

Contents

Introduction **1**

Load libraries **1**

Load data **1**
Data exclusion 2

Decision tree **3**

Mixed effects models **4**
Random effects structure 4
Fixed effects 6
Results 14
Summary 21

Plots **22**

Introduction

In this analysis, we leave out the first few trials.

Load libraries

```
library(lme4)
library(lattice)
library(gplots)
library(ggplot2)
library(sjPlot)
library(party)
library(Rmisc)
library(dplyr)
```

Function for converting from logit scale

```
logit2per = function(X){
  return(exp(X)/(1+exp(X)))
}
```

Load data

```
d = read.csv("../Data/Lab_Processed.csv")
```

Make *answer* a binary variable.

```
d$answer = d$answer=="Yes"
d$lastAnswer = d$lastAnswer=="Yes"
```

Relevel response phoneme and context.

```
d$responsePhoneme = relevel(d$responsePhoneme, 'other')
d$context = relevel(d$context, 'ST')
```

Center trial number, so that the intercept will reflect probabilities in the middle of the experiment.

```
d$trialNumber.center = d$trialNumber - 25
```

Data exclusion

We exclude participant 13 because they took much longer than other participants.

```
d = d[as.character(d$partID)!="13",]
```

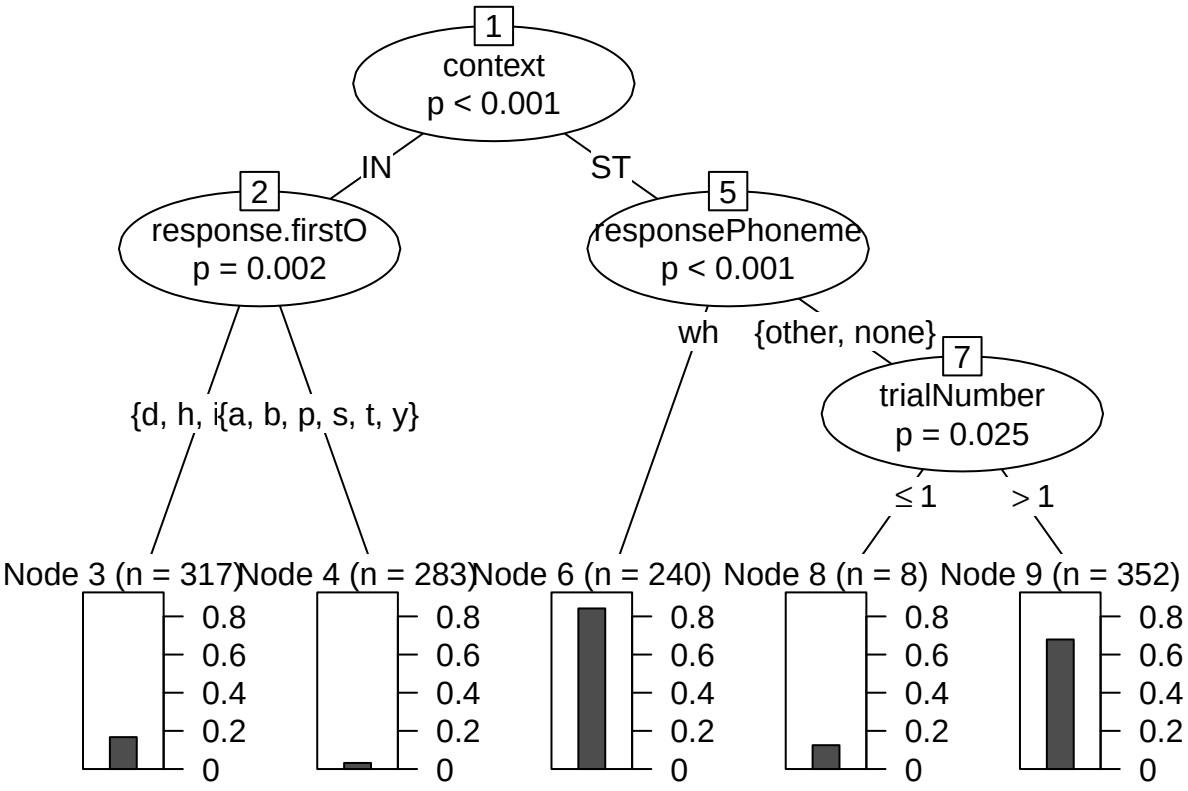
Decision tree

In order to get an idea of the structure of the data, we make a binary decision tree based on the data. We try to predict the participant's response by context, the first phoneme of the response, the type of turn the response was taken from and the participants' age, sex and the type of English they speak (American, British, or other, which are the main groups).

```
cx = ctree(answer ~
             context +responsePhoneme + responseType +
             Age + Sex + EnglishType +
             response.sex + context.sex +
             response.firstO + trialNumber + lastAnswer +
             blockName,
           data = d,
           controls = ctree_control(mincriterion = 0.95))
```

Plot the decision tree:

```
plot(cx, terminal_panel=node_barplot)
```



Context is the most important factor, followed by first phoneme of the response.

But we also see that there's a split by trial number. In this analysis we leave out the first 5 trials for each participant:

```
d = d[d$trialNumber >6,]
```

Mixed effects models

Make a series of mixed effects models. We can fix this using the "bobyqa" optimiser for both phases of the convergence and letting the algorithm run longer:

```
gcontrol = glmerControl(optimizer="bobyqa",optCtrl = list(maxfun=2e4))
```

(Note that several convergence algorithms were tested, and the three best fitting solutions had essentially no differences in fixed effect estimates)

Random effects structure

We have a good idea of what the random effects structure should be, but first we check whether there are significant differences by participant etc.

```
mA0 = glmer(
  answer ~ 1 +
    (1 | partID),
  data = d,
  family = binomial,
  control = gcontrol
)


mA0b = glmer(
  answer ~ 1 +
    (1 | blockName/partID) ,
  data = d,
  family = binomial,
  control = gcontrol
)
anova(mA0,mA0b)
```

```
## Data: d
## Models:
## mA0: answer ~ 1 + (1 | partID)
## mA0b: answer ~ 1 + (1 | blockName/partID)
##      Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## mA0   2 1478.2 1488.1 -737.08   1474.2
## mA0b  3 1479.6 1494.5 -736.80   1473.6 0.5692      1     0.4506
```

There is no significant improvement in the model when taking stimulus set into account. Because it complicates the analysis, we'll leave it out.

```
mA1 = glmer(
  answer ~ 1 +
    (1 | partID) +
    (1 | contextSample),
  data = d,
  family = binomial,
  control = gcontrol
)

mA2 = glmer(
  answer ~ 1 +
```

```
    (1 | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)

mA3 = glmer(
  answer ~ 1 +
    (1 + context| partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)

mA4 = glmer(
  answer ~ 1 +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)

anova(mA0, mA1, mA2, mA3, mA4)
```

```
## Data: d
## Models:
## mA0: answer ~ 1 + (1 | partID)
## mA1: answer ~ 1 + (1 | partID) + (1 | contextSample)
## mA2: answer ~ 1 + (1 | partID) + (1 | contextSample) + (1 | responseSample)
## mA3: answer ~ 1 + (1 + context | partID) + (1 | contextSample) + (1 |
## mA3:     responseSample)
## mA4: answer ~ 1 + (1 + context | partID) + (0 + responsePhoneme |
## mA4:     partID) + (1 | contextSample) + (1 | responseSample)
##     Df     AIC     BIC  logLik deviance   Chisq Chi Df Pr(>Chisq)
## mA0  2 1478.17 1488.14 -737.08  1474.17
## mA1  3  974.19  989.15 -484.10   968.19 505.975      1  < 2.2e-16 ***
## mA2  4  961.56  981.49 -476.78   953.56  14.637      1  0.0001303 ***
## mA3  6  922.95  952.85 -455.47   910.95  42.611      2  5.587e-10 ***
## mA4 12  921.52  981.34 -448.76   897.52  13.423      6  0.0367914 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

All proposed random effects significantly improve the fit of the model, except for the random slope for responsePhoneme by participant. This variable nearly doubles the number of model parameters, so we will leave it out.

Fixed effects

We are most interested in the effects of context and response type, but we need to check some other possible confounding variables.

Trial

```
m0 = glmer(
  answer ~  1 +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)

trial = glmer(
  answer ~ 1 + trialNumber.center +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)

trialQ = glmer(
  answer ~ 1 + trialNumber.center + I(trialNumber.center^2) +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unide
##  - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```
anova(m0,trial, trialQ)
```

```
## Data: d
## Models:
## m0: answer ~ 1 + (1 + context | partID) + (0 + responsePhoneme |
## m0:     partID) + (1 | contextSample) + (1 | responseSample)
## trial: answer ~ 1 + trialNumber.center + (1 + context | partID) + (0 +
## trial:     responsePhoneme | partID) + (1 | contextSample) + (1 | responseSample)
## trialQ: answer ~ 1 + trialNumber.center + I(trialNumber.center^2) + (1 +
## trialQ:     context | partID) + (0 + responsePhoneme | partID) + (1 |
## trialQ:     contextSample) + (1 | responseSample)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
```

```
## m0     12 887.11 946.65 -431.55   863.11
## trial  13 887.60 952.11 -430.80   861.60 1.5042      1     0.2200
## trialQ 14 889.57 959.05 -430.79   861.57 0.0270      1     0.8696
```

A significant effect of trial, but no significant quadratic term.

Previous answer

```
prevAns = glmer(
  answer ~ 1 + trialNumber.center + lastAnswer +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)
anova(trial,prevAns)
```

```
## Data: d
## Models:
## trial: answer ~ 1 + trialNumber.center + (1 + context | partID) + (0 +
## trial:     responsePhoneme | partID) + (1 | contextSample) + (1 | responseSample)
## prevAns: answer ~ 1 + trialNumber.center + lastAnswer + (1 + context |
## prevAns:     partID) + (0 + responsePhoneme | partID) + (1 | contextSample) +
## prevAns:     (1 | responseSample)
##         Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## trial   13 921.01 985.81 -447.51   895.01
## prevAns 14 922.12 991.91 -447.06   894.12 0.8887      1     0.3458
```

No significant effect of previous answer.

Sex of speakers in samples

```
contS = glmer(
  answer ~ 1 + trialNumber.center +
    context.sex +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)

respS = glmer(
  answer ~ 1 + trialNumber.center +
    context.sex + response.sex +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
```

```
)

contXrespS = glmer(
  answer ~ 1 + trialNumber.center +
    context.sex * response.sex +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)


anova(trial,contS, respS, contXrespS)
```

```
## Data: d
## Models:
## trial: answer ~ 1 + trialNumber.center + (1 + context | partID) + (0 +
## trial:     responsePhoneme | partID) + (1 | contextSample) + (1 | responseSample)
## contS: answer ~ 1 + trialNumber.center + context.sex + (1 + context |
## contS:     partID) + (0 + responsePhoneme | partID) + (1 | contextSample) +
## contS:     (1 | responseSample)
## respS: answer ~ 1 + trialNumber.center + context.sex + response.sex +
## respS:     (1 + context | partID) + (0 + responsePhoneme | partID) +
## respS:     (1 | contextSample) + (1 | responseSample)
## contXrespS: answer ~ 1 + trialNumber.center + context.sex * response.sex +
## contXrespS:     (1 + context | partID) + (0 + responsePhoneme | partID) +
## contXrespS:     (1 | contextSample) + (1 | responseSample)
##            Df    AIC     BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## trial      13 921.01  985.81 -447.51   895.01
## contS      14 922.36  992.14 -447.18   894.36 0.6557      1     0.4181
## respS      15 924.28  999.05 -447.14   894.28 0.0739      1     0.7858
## contXrespS 16 926.12 1005.87 -447.06   894.12 0.1667      1     0.6830
```

No significant effects of the sex of the speakers in the samples.

Sex of participants

```
sex = glmer(
  answer ~ 1 + trialNumber.center + Sex +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)

anova(trial,sex)
```

```
## Data: d
## Models:
```

```
## trial: answer ~ 1 + trialNumber.center + (1 + context | partID) + (0 +
## trial:     responsePhoneme | partID) + (1 | contextSample) + (1 | responseSample)
## sex: answer ~ 1 + trialNumber.center + Sex + (1 + context | partID) +
## sex:     (0 + responsePhoneme | partID) + (1 | contextSample) + (1 |
## sex:     responseSample)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## trial 13 921.01 985.81 -447.51   895.01
## sex   14 922.84 992.62 -447.42   894.84 0.1759      1     0.6749
```

No significant effect of the sex of the participant.

Type of English spoken

```
Etype = glmer(
  answer ~ 1 + trialNumber.center + EnglishType +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)

anova(trial,Etype)
```

```
## Data: d
## Models:
## trial: answer ~ 1 + trialNumber.center + (1 + context | partID) + (0 +
## trial:     responsePhoneme | partID) + (1 | contextSample) + (1 | responseSample)
## Etype: answer ~ 1 + trialNumber.center + EnglishType + (1 + context |
## Etype:     partID) + (0 + responsePhoneme | partID) + (1 | contextSample) +
## Etype:     (1 | responseSample)
##       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## trial 13 921.01 985.81 -447.51   895.01
## Etype 15 923.39 998.16 -446.70   893.39 1.6232      2     0.4441
```

No significant effec of the type of English the participant speaks.

Effects of Context and Response

The only significant confounding variable is trial.

```
context = glmer(
  answer ~ 1  + trialNumber.center +
    context +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
```

```
## $checkConv, : Model failed to converge with max|grad| = 0.0859039 (tol =
## 0.001, component 1)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unider
##  - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```
rPhon = glmer(
  answer ~ 1 + trialNumber.center +
    context + responsePhoneme +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)

conXrPh = glmer(
  answer ~ 1 + trialNumber.center +
    context * responsePhoneme +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model failed to conver

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unider
##  - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

```
anova(trial, context,rPhon, conXrPh)
```

```
## Data: d
## Models:
## trial: answer ~ 1 + trialNumber.center + (1 + context | partID) + (0 +
## trial:     responsePhoneme | partID) + (1 | contextSample) + (1 | responseSample)
## context: answer ~ 1 + trialNumber.center + context + (1 + context | partID) +
## context:     (0 + responsePhoneme | partID) + (1 | contextSample) + (1 |
## context:     responseSample)
## rPhon: answer ~ 1 + trialNumber.center + context + responsePhoneme +
## rPhon:     (1 + context | partID) + (0 + responsePhoneme | partID) +
## rPhon:     (1 | contextSample) + (1 | responseSample)
## conXrPh: answer ~ 1 + trialNumber.center + context * responsePhoneme +
## conXrPh:     (1 + context | partID) + (0 + responsePhoneme | partID) +
## conXrPh:     (1 | contextSample) + (1 | responseSample)
##         Df    AIC    BIC  logLik deviance   Chisq Chi Df Pr(>Chisq)
## trial   13 887.60 952.11 -430.80   861.60
## context 14 837.25 906.72 -404.62   809.25 52.3535      1  4.636e-13 ***
## rPhon   16 830.71 910.11 -399.36   798.71 10.5350      2   0.005156 **
```

```
## conXrPh 18 828.97 918.29 -396.48   792.97  5.7458      2   0.056536 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction between Sex and responses

```
Sex = glmer(
  answer ~ 1  + trialNumber.center +
    context * responsePhoneme +
    Sex +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample),
  data = d,
  family = binomial,
  control = gcontrol
)
```

```
## Warning in optwrap(optimizer, devfun, start, rho$lower, control =
## control, : convergence code 1 from bobyqa: bobyqa -- maximum number of
## function evaluations exceeded
```

```
SexXresp = glmer(
  answer ~ 1  + trialNumber.center +
    context * responsePhoneme +
    Sex*responsePhoneme +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)
```

```
## Warning in optwrap(optimizer, devfun, start, rho$lower, control =
## control, : convergence code 1 from bobyqa: bobyqa -- maximum number of
## function evaluations exceeded
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.00964715 (tol =
## 0.001, component 1)
```

```
SexXcon = glmer(
  answer ~ 1 + trialNumber.center +
    context * responsePhoneme +
    Sex*responsePhoneme +
    Sex:context +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)
```

```
## Warning in optwrap(optimizer, devfun, start, rho$lower, control =
## control, : convergence code 1 from bobyqa: bobyqa -- maximum number of
## function evaluations exceeded
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.0269849 (tol =
## 0.001, component 1)
```

```
SxXcoXre = glmer(
  answer ~ 1 + trialNumber.center +
    context * responsePhoneme +
    Sex*responsePhoneme*context +
    (1 + context | partID) +
    (0 + responsePhoneme | partID) +
    (1 | contextSample) +
    (1 | responseSample) ,
  data = d,
  family = binomial,
  control = gcontrol
)
```

```
## Warning in optwrap(optimizer, devfun, start, rho$lower, control =
## control, : convergence code 1 from bobyqa: bobyqa -- maximum number of
## function evaluations exceeded
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.108048 (tol =
## 0.001, component 1)
```

```
anova(conXrPh, Sex, SexXresp, SexXcon, SxXcoXre)
```

```
## Data: d
## Models:
## conXrPh: answer ~ 1 + trialNumber.center + context * responsePhoneme +
## conXrPh:     (1 + context | partID) + (0 + responsePhoneme | partID) +
## conXrPh:     (1 | contextSample) + (1 | responseSample)
## Sex: answer ~ 1 + trialNumber.center + context * responsePhoneme +
## Sex:     Sex + (1 + context | partID) + (0 + responsePhoneme | partID) +
## Sex:     (1 | contextSample) + (1 | responseSample)
## SexXresp: answer ~ 1 + trialNumber.center + context * responsePhoneme +
## SexXresp:     Sex * responsePhoneme + (1 + context | partID) + (0 + responsePhoneme |
## SexXresp:     partID) + (1 | contextSample) + (1 | responseSample)
## SexXcon: answer ~ 1 + trialNumber.center + context * responsePhoneme +
## SexXcon:     Sex * responsePhoneme + Sex:context + (1 + context | partID) +
## SexXcon:     (0 + responsePhoneme | partID) + (1 | contextSample) + (1 |
## SexXcon:     responseSample)
## SxXcoXre: answer ~ 1 + trialNumber.center + context * responsePhoneme +
## SxXcoXre:     Sex * responsePhoneme * context + (1 + context | partID) +
## SxXcoXre:     (0 + responsePhoneme | partID) + (1 | contextSample) + (1 |
## SxXcoXre:     responseSample)
##          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## conXrPh  18 828.97 918.29 -396.48   792.97
## Sex      19 830.20 924.48 -396.10   792.20 0.7651      1     0.3817
## SexXresp 21 832.72 936.93 -395.36   790.72 1.4813      2     0.4768
## SexXcon  22 833.89 943.06 -394.95   789.89 0.8288      1     0.3626
## SxXcoXre 24 836.85 955.94 -394.42   788.85 1.0420      2     0.5939
```

No effect by sex of participant.

Results

Model comparison

```
anova(trial, context,rPhon, conXrPh)
```

```
## Data: d
## Models:
## trial: answer ~ 1 + trialNumber.center + (1 + context | partID) + (0 +
## trial:     responsePhoneme | partID) + (1 | contextSample) + (1 | responseSample)
## context: answer ~ 1 + trialNumber.center + context + (1 + context | partID) +
## context:     (0 + responsePhoneme | partID) + (1 | contextSample) + (1 |
## context:     responseSample)
## rPhon: answer ~ 1 + trialNumber.center + context + responsePhoneme +
## rPhon:     (1 + context | partID) + (0 + responsePhoneme | partID) +
## rPhon:     (1 | contextSample) + (1 | responseSample)
## conXrPh: answer ~ 1 + trialNumber.center + context * responsePhoneme +
## conXrPh:     (1 + context | partID) + (0 + responsePhoneme | partID) +
## conXrPh:     (1 | contextSample) + (1 | responseSample)
##         Df    AIC    BIC  logLik deviance   Chisq Chi Df Pr(>Chisq)
## trial   13 887.60 952.11 -430.80   861.60
## context 14 837.25 906.72 -404.62   809.25 52.3535      1  4.636e-13 ***
## rPhon   16 830.71 910.11 -399.36   798.71 10.5350      2   0.005156 **
## conXrPh 18 828.97 918.29 -396.48   792.97  5.7458      2   0.056536 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Fixed effects

Model estimates:

```
finalModel = conXrPh
save(finalModel, file="../results/FinalModel.Rdat")
summary(finalModel)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: answer ~ 1 + trialNumber.center + context * responsePhoneme +
##     (1 + context | partID) + (0 + responsePhoneme | partID) +
##     (1 | contextSample) + (1 | responseSample)
##    Data: d
## Control: gcontrol
##
##      AIC      BIC   logLik deviance df.resid
##    829.0    918.3   -396.5    793.0     1038
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.3893 -0.2605 -0.1137  0.3839  6.0588
##
## Random effects:
##  Groups         Name                 Variance Std.Dev. Corr
##  responseSample (Intercept)          0.1053   0.3245
##  contextSample  (Intercept)          1.3162   1.1473
##  partID         responsePhonemeother 0.3115   0.5581
##                 responsePhonemenone  1.9539   1.3978    0.85
##                 responsePhonemewh    0.5574   0.7466   -0.25  0.30
##  partID.1       (Intercept)          0.4371   0.6611
##                 contextIN            1.7389   1.3187   -1.00
## Number of obs: 1056, groups:
## responseSample, 51; contextSample, 50; partID, 24
##
## Fixed effects:
##                                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    0.9248626  0.0008482    1090   <2e-16 ***
## trialNumber.center             0.0110693  0.0008434      13   <2e-16 ***
## contextIN                     -4.1934096  0.0008481   -4944   <2e-16 ***
## responsePhonemenone            0.2934238  0.0008478     346   <2e-16 ***
## responsePhonemewh              1.1856380  0.0008479    1398   <2e-16 ***
## contextIN:responsePhonemenone -1.7141119  0.0008479   -2022   <2e-16 ***
## contextIN:responsePhonemewh   -0.4175329  0.0008479    -492   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                    (Intr) trlNm. cntxIN rspnsPhnmn rspnsPhnmw
## trlNmbr.cnt        0.000
## contextIN          0.000  0.000
## rspnsPhnmnn        0.000  0.000  0.000
## rspnsPhnmwh        0.000  0.000  0.000  0.000
## cntxtIN:rspnsPhnmn 0.000  0.000  0.000  0.000      0.000
```

```
## cntxtIN:rspnsPhnmw 0.000  0.000  0.000  0.000      0.000
##                    cntxtIN:rspnsPhnmn
## trlNmbr.cnt
## contextIN
## rspnsPhnmnn
## rspnsPhnmwh
## cntxtIN:rspnsPhnmn
## cntxtIN:rspnsPhnmw 0.000
## convergence code: 0
## Model failed to converge with max|grad| = 0.0846642 (tol = 0.001, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
## Model is nearly unidentifiable: large eigenvalue ratio
##  - Rescale variables?
```

Relevel the response phoneme to see other comparisons:

```
d2 = d
d2$responsePhoneme = relevel(d2$responsePhoneme,"wh")
fm2 = update(finalModel, data=d2)
```

```
## Warning in optwrap(optimizer, devfun, start, rho$lower, control =
## control, : convergence code 1 from bobyqa: bobyqa -- maximum number of
## function evaluations exceeded
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control
## $checkConv, : Model failed to converge with max|grad| = 0.00319295 (tol =
## 0.001, component 1)
```

```
summary(fm2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: answer ~ 1 + trialNumber.center + context * responsePhoneme +
##     (1 + context | partID) + (0 + responsePhoneme | partID) +
##     (1 | contextSample) + (1 | responseSample)
##    Data: d2
## Control: gcontrol
##
##      AIC      BIC   logLik deviance df.resid
##    828.2    917.5   -396.1    792.2     1038
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4825 -0.2505 -0.1020  0.3725  6.3086
##
## Random effects:
##  Groups         Name                Variance Std.Dev. Corr
##  responseSample (Intercept)         0.1398   0.3739
##  contextSample  (Intercept)         1.3701   1.1705
##  partID         responsePhonemewh   0.6381   0.7988
##                 responsePhonemeother 0.3266   0.5715   -0.21
##                 responsePhonemenone  2.3098   1.5198    0.33  0.85
##  partID.1       (Intercept)         0.4515   0.6720
##                 contextIN           1.8328   1.3538   -1.00
```

```
## Number of obs: 1056, groups:
## responseSample, 51; contextSample, 50; partID, 24
##
## Fixed effects:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    2.326483   0.422939   5.501 3.78e-08 ***
## trialNumber.center             0.011697   0.008282   1.412  0.15788
## contextIN                     -4.986142   0.622191  -8.014 1.11e-15 ***
## responsePhonemeother          -1.286182   0.394752  -3.258  0.00112 **
## responsePhonemenone           -0.921877   0.645271  -1.429  0.15310
## contextIN:responsePhonemeother  0.495939   0.532241   0.932  0.35144
## contextIN:responsePhonemenone  -1.667247   0.962515  -1.732  0.08324 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##                    (Intr) trlNm. cntxIN rspnsPhnmt rspnsPhnmn
## trlNmbr.cnt        -0.005
## contextIN          -0.686 -0.076
## rspnsPhnmth        -0.616 -0.081  0.320
## rspnsPhnmnn        -0.306 -0.018  0.188  0.455
## cntxtIN:rspnsPhnmt  0.302  0.069 -0.394 -0.478     -0.209
## cntxtIN:rspnsPhnmn  0.182 -0.021 -0.211 -0.185     -0.389
##                    cntxtIN:rspnsPhnmt
## trlNmbr.cnt
## contextIN
## rspnsPhnmth
## rspnsPhnmnn
## cntxtIN:rspnsPhnmt
## cntxtIN:rspnsPhnmn  0.314
## convergence code: 1
## Model failed to converge with max|grad| = 0.00319295 (tol = 0.001, component 1)
```

Confidence intervals (through Wald method):

```
CI = confint(finalModel,parm="beta_", method="Wald")
cx = summary(finalModel)$coef
cx = cbind(cx[,1],CI,cx[,2:4])
cx2 = cx
for(i in 1:5){cx2[,i] = round(cx2[,i],3)}
cx2
```

```
##                                    2.5 % 97.5 % Std. Error   z value
## (Intercept)                    0.925  0.923  0.927      0.001  1090.419
## trialNumber.center             0.011  0.009  0.013      0.001    13.124
## contextIN                     -4.193 -4.195 -4.192      0.001 -4944.200
## responsePhonemenone            0.293  0.292  0.295      0.001   346.094
## responsePhonemewh              1.186  1.184  1.187      0.001  1398.377
## contextIN:responsePhonemenone -1.714 -1.716 -1.712      0.001 -2021.629
## contextIN:responsePhonemewh   -0.418 -0.419 -0.416      0.001  -492.407
##                                   Pr(>|z|)
## (Intercept)                   0.000000e+00
## trialNumber.center            2.390544e-39
## contextIN                     0.000000e+00
## responsePhonemenone           0.000000e+00
```

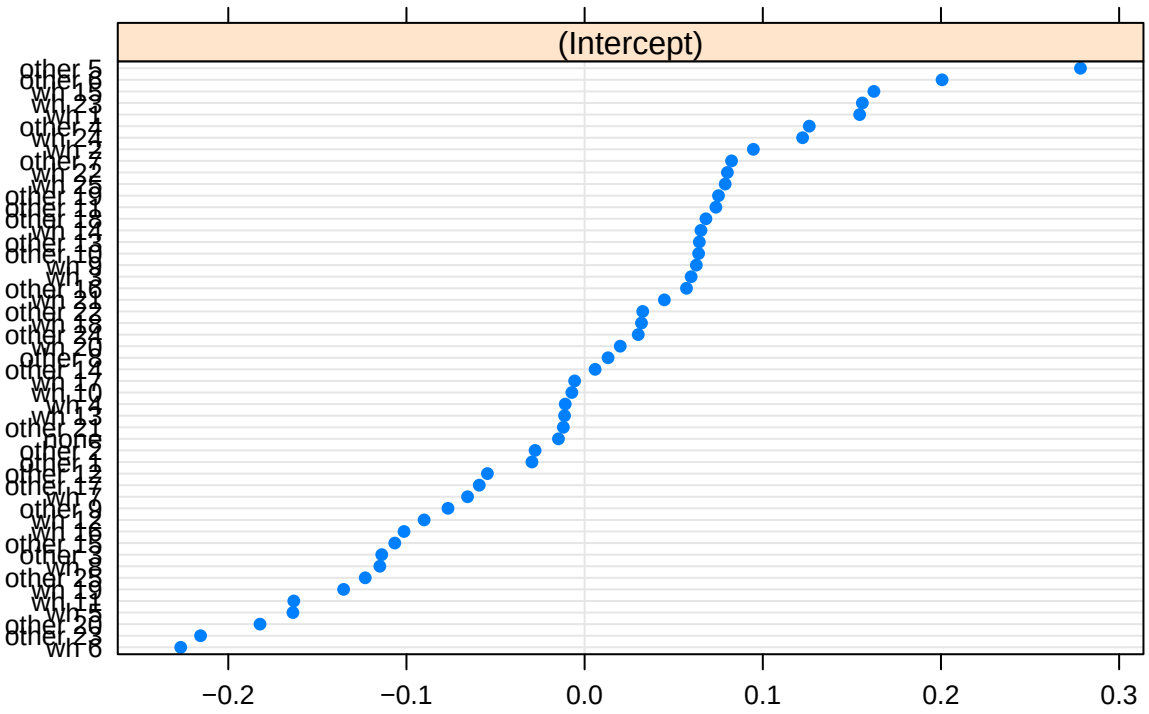
```
## responsePhonemewh              0.000000e+00
## contextIN:responsePhonemenone 0.000000e+00
## contextIN:responsePhonemewh   0.000000e+00
```

```
write.csv(cx, "../results/FinalModelCoefficients.csv")
```

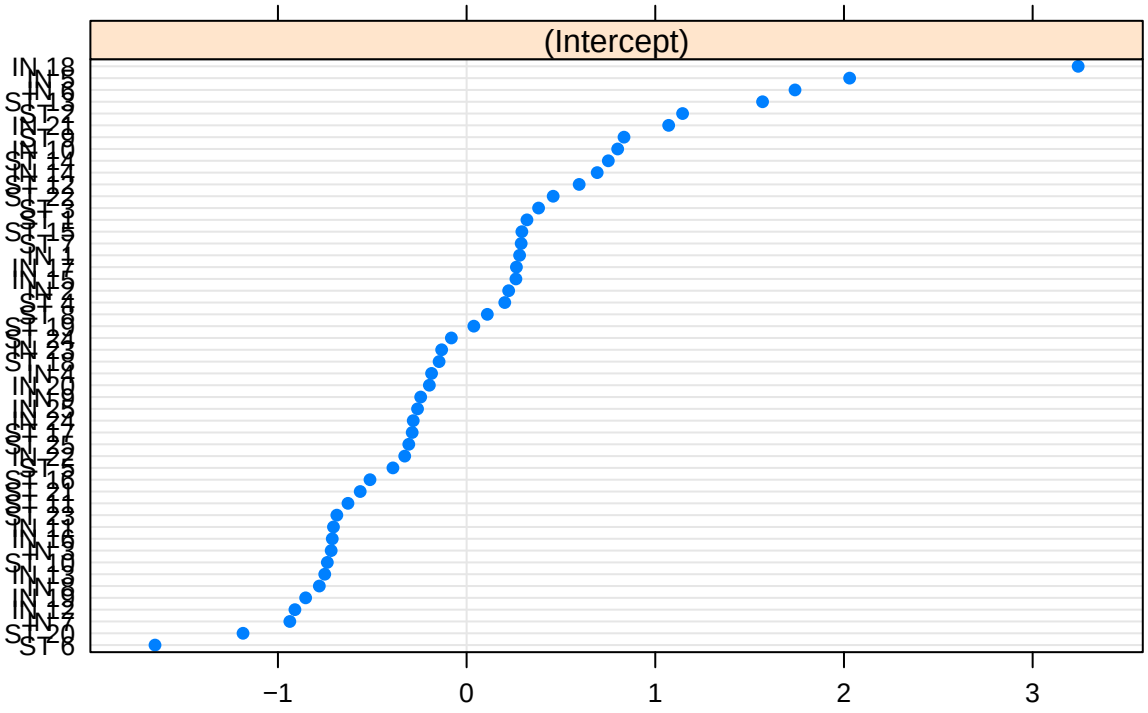
Random effects

```
dotplot(ranef(finalModel))
```

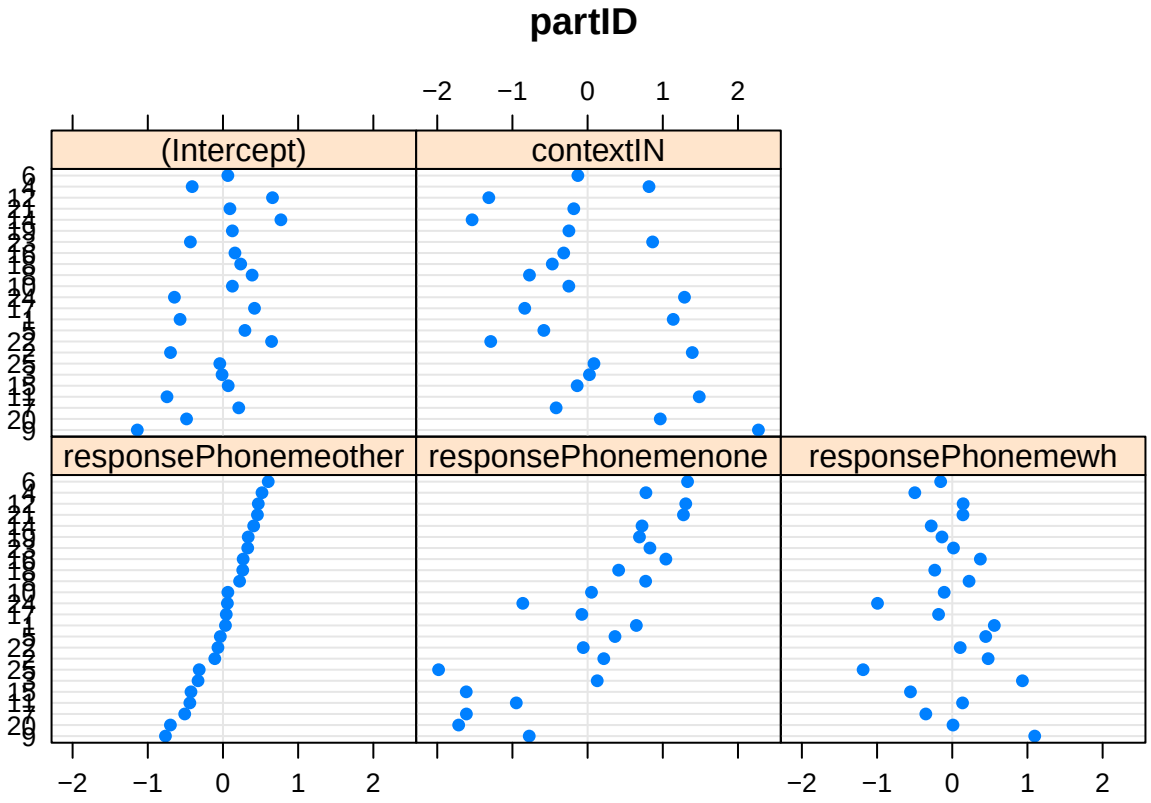
```
## $responseSample
```




```
##
## $contextSample
```


contextSample
(Intercept)
−1
0
1
2
3

```
##
## $partID
```


partID
(Intercept)
contextIN
responsePhonemeother
responsePhonemenone
responsePhonemewh
−2
−1
0
1
2

Summary

Here is a summary of the main results:

There was a significant main effect of context (log likelihood difference = 27 , df = 2 , Chi Squared = 53.86 , p = 2e-12).

There was a significant main effect of phoneme (log likelihood difference = 5.3 , df = 2 , Chi Squared = 10.54 , p = 0.0052).

There was a marginal interaction between context and phoneme (log likelihood difference = 2.9 , df = 2 , Chi Squared = 5.75 , p = 0.057).

Work out model esimates for probabilities in each condition:

```
# prob of responding 'yes' when:
# Context = ST, other response
logit2per(fixef(finalModel)[1])[[1]]
```

```
## [1] 0.7160319
```

```
# Context = ST, no response
logit2per(fixef(finalModel)[1] + fixef(finalModel)["responsePhonemenone"]  )[[1]]
```

```
## [1] 0.7717618
```

```
# Context = ST, wh
logit2per(fixef(finalModel)[1] + fixef(finalModel)["responsePhonemewh"]  )[[1]]
```

```
## [1] 0.8919196
```

```
# Context = IN, other response
logit2per(fixef(finalModel)[1] + fixef(finalModel)["contextIN"])[[1]]
```

```
## [1] 0.03666612
```

```
# Context = IN, no response
logit2per(fixef(finalModel)[1] +
            fixef(finalModel)["contextIN"] +
            fixef(finalModel)["responsePhonemenone"])[[1]]
```

```
## [1] 0.04856246
```

```
# Context = IN, wh
logit2per(fixef(finalModel)[1] +
            fixef(finalModel)["contextIN"] +
            fixef(finalModel)["responsePhonemewh"])[[1]]
```

```
## [1] 0.1107691
```

Plots

Fixed effects estimates:

```
feLabels = matrix(c(
  "(Intercept)"              ,"Intercept"      , NA,
  "trialNumber.center", "Trial",NA,
  "contextST", "Context = Statement", "context",
  "contextIN", "Context = Initiating", "context",
  "responsePhonemenone", "no response", 'rPhon',
  "responsePhonemewh", "wh response", 'rPhon',
  "contextIN:responsePhonemenone", "Context = In: no response", "conXrPh",
  "contextIN:responsePhonemewh", "Context = In: wh response", "conXrPh"
), ncol=3, byrow = T)

feLabels2 = as.vector(feLabels[match(names(fixef(finalModel)),feLabels[,1]),2])
```

```
sjp.glmer(finalModel,'fe' ,
          show.intercept = T,
          geom.colors = c(1,1),
          axis.title = "Odds of selecting question",
          y.offset = 0.2
)
```

```
## Warning: Deprecated, use tibble::rownames_to_column() instead.
```