

Analysis of effects of cortical inactivation on vowel discrimination in noise using logistic regression

Version History:

- Created: 2021-09-09 Stephen Town
- Ported to notebook: 2022-05-19

Combine data across ferrets and run a generalized linear mixed model to determine the effect of cooling on log-odds of correct response, with subject and test session as random effects, and cooling as fixed effects

The individual observations are at the trial level, which is important for analysis of error magnitude, but means that the data for trial accuracy (correct/error) is in binary format.

```
rm(list = ls())

suppressPackageStartupMessages( library(dplyr))
suppressPackageStartupMessages( library(tidyverse))

suppressPackageStartupMessages( library(lme4))
suppressPackageStartupMessages( library(DHARMA))
suppressPackageStartupMessages( library(scales))
suppressPackageStartupMessages( library(sjPlot))
```

1. Load data

Load data from ferrets with working cooling loops / optic fibers and select specific columns.

```
data_path = '/home/stephen/Github/Vowel_Discrimination_In_Noise/Results/Vowels_in_Noise/data/analysis'

magnum = read_csv( file.path( data_path, 'F1311.csv'), show_col_types = FALSE)
robin  = read_csv( file.path( data_path, 'F1509.csv'), show_col_types = FALSE)
mimi   = read_csv( file.path( data_path, 'F1706.csv'), show_col_types = FALSE)

cols = c('fNum','treatment',"vowel","Mask","Correct","vowel_level","SNR","CenterReward")
df = rbind(magnum[cols], robin[cols], mimi[cols])

df$fNum <- as.factor(df$fNum)
head(df)
```

```
## # A tibble: 6 x 8
##   fNum treatment vowel Mask      Correct vowel_level SNR CenterReward
##   <fct> <lg1>    <chr> <chr>    <dbl>    <dbl> <dbl>    <dbl>
## 1 1311 FALSE    a     Restricted 0        66    -4        0
## 2 1311 FALSE    a     Clean      1        66    -4        0
## 3 1311 FALSE    i     Clean      1        66    -4        0
## 4 1311 FALSE    a     Restricted 1        66    -4        0
## 5 1311 FALSE    a     Restricted 0        66    -4        1
## 6 1311 FALSE    i     Restricted 0        66    -4        1
```

```
table(df$fNum)
```

```
##  
## 1311 1509 1706  
## 1914  603  352
```

```
nrow(df)
```

```
## [1] 2869
```

2. Transform

Although not strictly necessary for models without regularization, it is helpful for the interpretation of coefficients if we scale the sound level of vowels between 0 and 1. We then count the number of correct trials in each unique condition, and the total number of trials; again this isn't strictly needed, but count data can be more convenient to work with when inspecting model outputs.

```
df$vowel_level_orig <- df$vowel_level  # Save a copy for reconstruction  
df$vowel_level <- rescale(df$vowel_level)
```

```
counts = df %>%  
  group_by(fNum, treatment, Mask, vowel, vowel_level, CenterReward) %>%  
  summarise(  
    nCorrect = sum(Correct),  
    nTotal = n(),  
    .groups = 'keep'  
  ) %>%  
  mutate(pCorrect = nCorrect / nTotal * 100)
```

```
head(counts)
```

```
## # A tibble: 6 x 9  
## # Groups:   fNum, treatment, Mask, vowel, vowel_level, CenterReward [6]  
##   fNum treatment Mask vowel vowel_level CenterReward nCorrect nTotal pCorrect  
##   <fct> <lgl>    <chr> <chr>      <dbl>         <dbl>    <dbl> <int>    <dbl>  
## 1 1311 FALSE    Clean a      0.591          0      41     58     70.7  
## 2 1311 FALSE    Clean a      0.591          1      17     24     70.8  
## 3 1311 FALSE    Clean a      0.727          0      59     77     76.6  
## 4 1311 FALSE    Clean a      0.727          1      18     20      90  
## 5 1311 FALSE    Clean a      1            0      58     75     77.3  
## 6 1311 FALSE    Clean a      1            1      23     27     85.2
```

3. Fit Model

Fit a logistic regression model with treatment (on/off), mask (clean or with noise) and vowel sound level as factors

```
mdl_1 = glmer(  
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment*Mask + vowel_level + CenterReward + (1|fNum),  
  family=binomial,  
  data=counts)
```

```
summary(mdl_1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )
```

```
## Formula:
## cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment * Mask + vowel_level +
##   CenterReward + (1 | fNum)
##   Data: counts
##
##      AIC      BIC   logLik deviance df.resid
##    696.0    717.4   -341.0    682.0     150
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.0176 -0.9360  0.4064  0.9401  3.1355
##
## Random effects:
##   Groups Name      Variance Std.Dev.
##   fNum   (Intercept) 0.04494  0.212
## Number of obs: 157, groups: fNum, 3
##
## Fixed effects:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.1623     0.1913   6.078 1.22e-09 ***
## treatmentTRUE      0.1886     0.1691   1.115  0.26468
## MaskRestricted    -0.5322     0.1138  -4.676 2.92e-06 ***
## vowel_level       0.5561     0.2107   2.639  0.00831 **
## CenterReward     -0.0857     0.1045  -0.820  0.41220
## treatmentTRUE:MaskRestricted -0.6459     0.1977  -3.267  0.00109 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) trTRUE MskRst vwl_lv CntrRw
## tretmntTRUE -0.286
## MaskRstrctd -0.402  0.426
## vowel_level -0.541  0.044 -0.022
## CenterRewrd -0.050 -0.040 -0.042  0.009
## trtmTRUE:MR  0.226 -0.852 -0.545 -0.039  0.023
```

3.2. Two two-way interactions

Add interaction between mask and vowel level (predicted)

```
mdl_2 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment*Mask + Mask*vowel_level + CenterReward + (1|fNum),
  family=binomial,
  data=counts)

summary(mdl_2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##   Family: binomial ( logit )
## Formula: cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment * Mask + Mask *
##   vowel_level + CenterReward + (1 | fNum)
##   Data: counts
##
##      AIC      BIC   logLik deviance df.resid
```

```

##      694.8      719.3     -339.4      678.8      149
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.8317 -0.8656  0.3831  0.9676  3.1415
##
## Random effects:
##   Groups Name            Variance Std.Dev.
##   fNum    (Intercept) 0.05201  0.2281
## Number of obs: 157, groups: fNum, 3
##
## Fixed effects:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   1.47391    0.26616   5.538 3.06e-08 ***
## treatmentTRUE                   0.15119    0.17022   0.888 0.374408
## MaskRestricted                 -0.95166    0.26329  -3.614 0.000301 ***
## vowel_level                     0.11199    0.32526   0.344 0.730617
## CenterReward                  -0.07997    0.10437  -0.766 0.443558
## treatmentTRUE:MaskRestricted -0.60313    0.19909  -3.029 0.002450 **
## MaskRestricted:vowel_level     0.61257    0.34405   1.780 0.075006 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) trTRUE MskRst vwl_lv CntrRw tTRUE:
## tretmntTRUE -0.283
## MaskRstrctd -0.731  0.292
## vowel_level -0.765  0.119  0.691
## CenterRewrd -0.013 -0.046 -0.046 -0.022
## trtmTRUE:MR  0.238 -0.854 -0.340 -0.113  0.029
## MskRstrct:_  0.666 -0.120 -0.902 -0.766  0.031  0.117

```

3.3. Model 3: Three-way Interaction

Add full combination of interactions between mask, treatment and level (not sure why this would be interesting *a priori*). Note that the AIC resulting from this model is higher than model

```

mdl_3 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment*Mask*vowel_level + CenterReward + (1|fNum),
  family=binomial,
  data=counts)

summary(mdl_3)

```

```

## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##   Family: binomial ( logit )
##   Formula:
##   cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment * Mask * vowel_level +
##     CenterReward + (1 | fNum)
##   Data: counts
##
##      AIC      BIC    logLik deviance df.resid
##    697.0    727.6   -338.5    677.0      147
##

```

```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.7244 -0.8781  0.4125  0.9469  3.1423
##
## Random effects:
##   Groups Name      Variance Std.Dev.
##   fNum    (Intercept) 0.05225  0.2286
## Number of obs: 157, groups: fNum, 3
##
## Fixed effects:
##                                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)                       1.51034     0.30138   5.011  5.4e-07
## treatmentTRUE                      0.05093     0.41373   0.123  0.902027
## MaskRestricted                     -1.12212     0.31530  -3.559  0.000372
## vowel_level                        0.06183     0.38217   0.162  0.871468
## CenterReward                      -0.08410     0.10458  -0.804  0.421289
## treatmentTRUE:MaskRestricted        -0.20672     0.48238  -0.429  0.668253
## treatmentTRUE:vowel_level          0.15953     0.60005   0.266  0.790346
## MaskRestricted:vowel_level          0.87203     0.42958   2.030  0.042362
## treatmentTRUE:MaskRestricted:vowel_level -0.63444     0.70091  -0.905  0.365380
##
## (Intercept)                       ***
## treatmentTRUE
## MaskRestricted                       ***
## vowel_level
## CenterReward
## treatmentTRUE:MaskRestricted
## treatmentTRUE:vowel_level
## MaskRestricted:vowel_level          *
## treatmentTRUE:MaskRestricted:vowel_level
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr) trTRUE MskRst vwl_lv CntrRw trTRUE:MR tTRUE:_ MskR:_
## tretmntTRUE -0.530
## MaskRstrctd -0.749  0.511
## vowel_level -0.821  0.523  0.726
## CenterRewrd -0.030  0.017 -0.014  0.001
## trtmTRUE:MR  0.453 -0.858 -0.618 -0.450 -0.029
## trtmnTRUE:_  0.464 -0.911 -0.447 -0.524 -0.039  0.782
## MskRstrct:_  0.693 -0.476 -0.932 -0.770 -0.001  0.582    0.475
## trTRUE:MR:_ -0.397  0.780  0.548  0.445  0.045 -0.911    -0.857  -0.595
```

4. Model Selection

From the three variants we looked at, the AIC criteria would favor model 2 (two two-way interactions), and this seems reasonable, given we would expect both the interaction between mask and sound level, and mask and treatment to be important.

```
anova mdl_1, mdl_2, mdl_3)
```

```
## Data: counts
```

```
## Models:
```

```
## mdl_1: cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment * Mask + vowel_level + CenterReward + (1 |
```

```
## mdl_2: cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment * Mask + Mask * vowel_level + CenterReward
## mdl_3: cbind(nCorrect, nTotal - nCorrect) ~ 1 + treatment * Mask * vowel_level + CenterReward + (1 |
##      npar      AIC      BIC logLik deviance  Chisq Df Pr(>Chisq)
## mdl_1      7 695.99 717.38 -341.0   681.99
## mdl_2      8 694.80 719.25 -339.4   678.80 3.1877  1    0.07419 .
## mdl_3     10 697.01 727.57 -338.5   677.01 1.7957  2    0.40744
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#tab_model(mdl_2)
```

5. Check model

5.1. Predictions

Use the model to predict the correct response probability. Then convert to correct trial numbers for each unique condition and aggregate over covariates (center reward, vowel level) for plotting.

```
counts$fit <- predict(mdl_2, counts, type="response")

# Count trials for ferret, treatment, mask and sound level
agg_by_level <- counts %>%
  mutate(
    predicted_correct = fit * nTotal
  ) %>%
  group_by(
    fNum, treatment, Mask, vowel_level
  ) %>%
  summarise(
    nCorrect = sum(nCorrect),
    nTotal = sum(nTotal),
    predicted_correct = sum(predicted_correct),
    .groups = 'keep'
  ) %>%
  mutate(
    pCorrect = nCorrect / nTotal * 100,
    fit_p = predicted_correct / nTotal * 100
  )

# Restore original vowel levels for plotting
agg_by_level$vowel_level <- rescale(
  agg_by_level$vowel_level,
  to = c(min(df$vowel_level_orig) , max(df$vowel_level_orig)))

# Calculate SNR
agg_by_level$SNR = agg_by_level$vowel_level - 70
```

Plot with vowel level

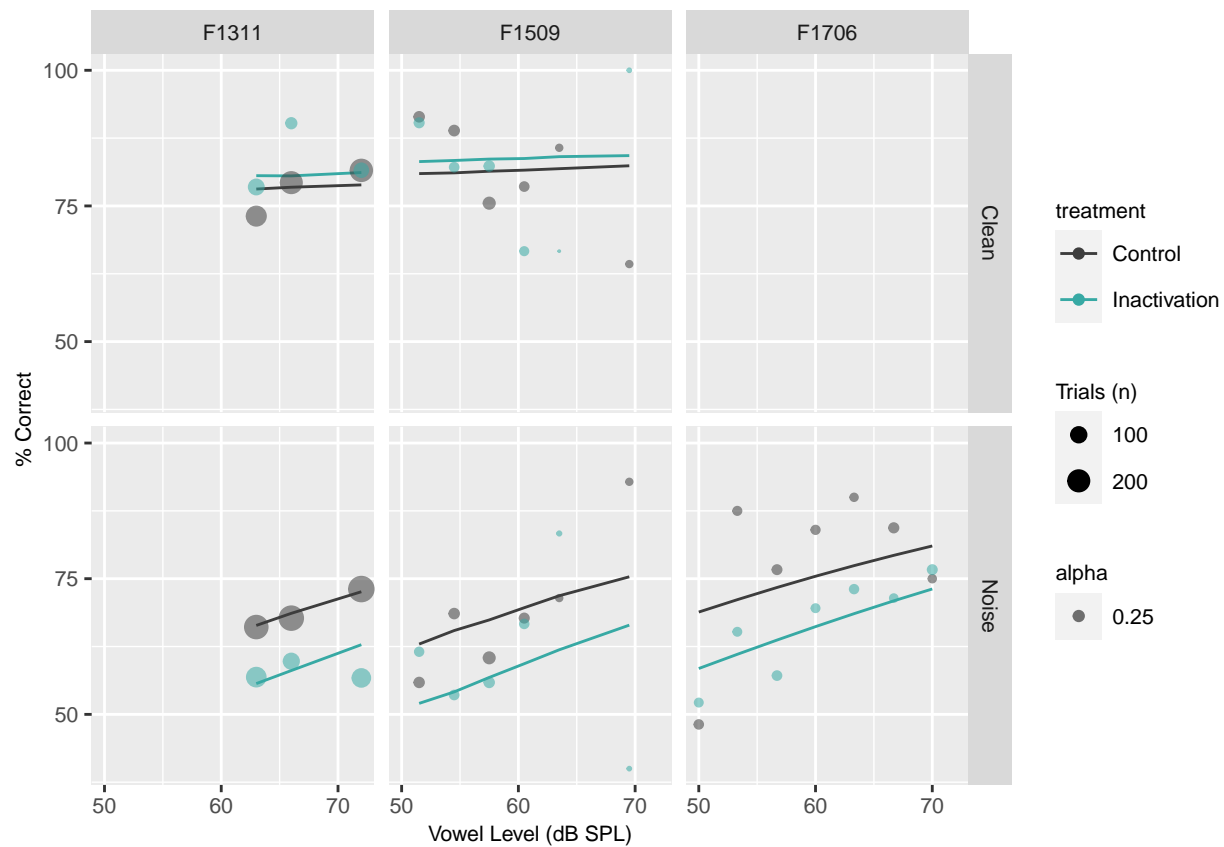
```
# Change name of clean trials to alter order of the rows in ggplot facets
agg_by_level$Mask[agg_by_level$Mask == 'xClean'] <- 'Clean'

ggplot(
  agg_by_level,
  aes(
```

```

    x = vowel_level,
    color = treatment
  )
) +
geom_line(
  aes(y = fit_p)
) +
geom_point(
  shape = 19,
  aes(
    y = pCorrect,
    size = nTotal,
    alpha = 0.25
  )
) +
scale_size(range = c(0, 4)) +
labs(
  x = "Vowel Level (dB SPL)",
  y = "% Correct",
  size='Trials (n)'
) +
scale_color_manual(
  labels = c("Control", "Inactivation"),
  values = c("#3d3d3d", "#37a9a4")
) +
scale_x_continuous(breaks=seq(50,70,by=10)) +
scale_y_continuous(breaks=seq(50,100,by=25)) +
theme(
  axis.text = element_text(size = 7.5),
  axis.title = element_text(size = 8),
  strip.text = element_text(size = 8),
  legend.title = element_text(size=8),
  legend.text = element_text(size=7.5)
  #strip.background = element_rect(colour="black", fill="white")
) +
facet_grid(
  Mask ~ fNum,
  labeller = labeller(
    Mask = c("Restricted" = "Noise", "Clean" = "Clean"),
    fNum = c("1311" = "F1311", "1509" = "F1509", "1706" = "F1706")
  )
)

```



```
ggsave(filename = "Noise_cooling_predictions.png", width = 12, height = 6.5, units = "cm", dpi=300)
```

5.2. Model Checking

```
simulationOutput <- simulateResiduals(fittedModel = mdl_2, n=500, plot = T)
```


DHARMA residual

