R Code for Human Listeners' Perception of Behavioural Context and Core Affect Dimensions in Chimpanzee Vocalisations

Roza Kamiloglu r.g.kamiloglu@uva.nl

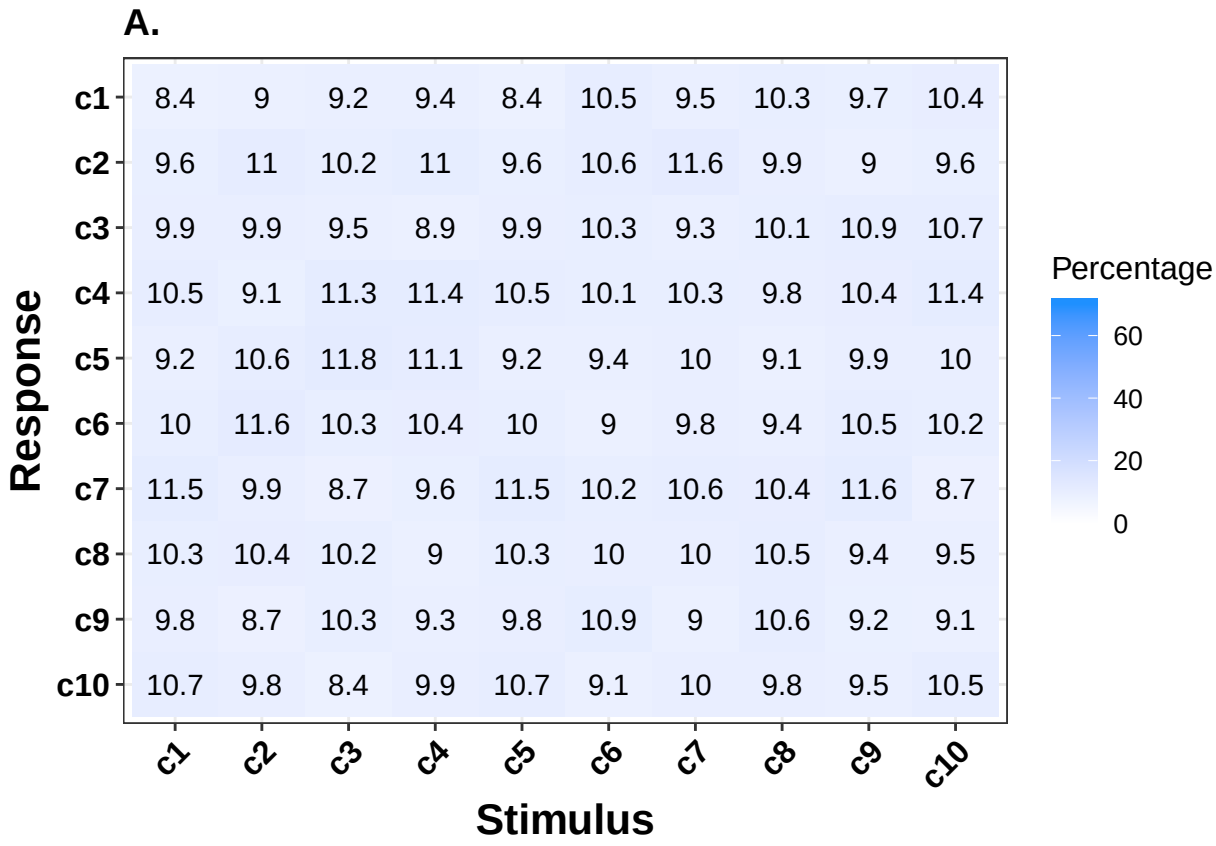
Explanation of the study and the code

This code: 1) Produces confusion matrices for accurate perception of each behavioral context, arousal level and valence based on the categoriation task (Experiment 1). 2) Displays d primes for the match-to-context task (Experiment 2). 3) Implements GLMM to test which acoustic parameters of the vocalizations predict human' ability to accurately perceive behavioral context (context-matching task), arousal and valence in chimpanzee vocalizations. Predictors are SCoG, Duration, MeanF0, StdDevF0, Mean HNR and Max HNR, DV is recognition (binary), and random factors are Participant ID and Chimpanzee ID.

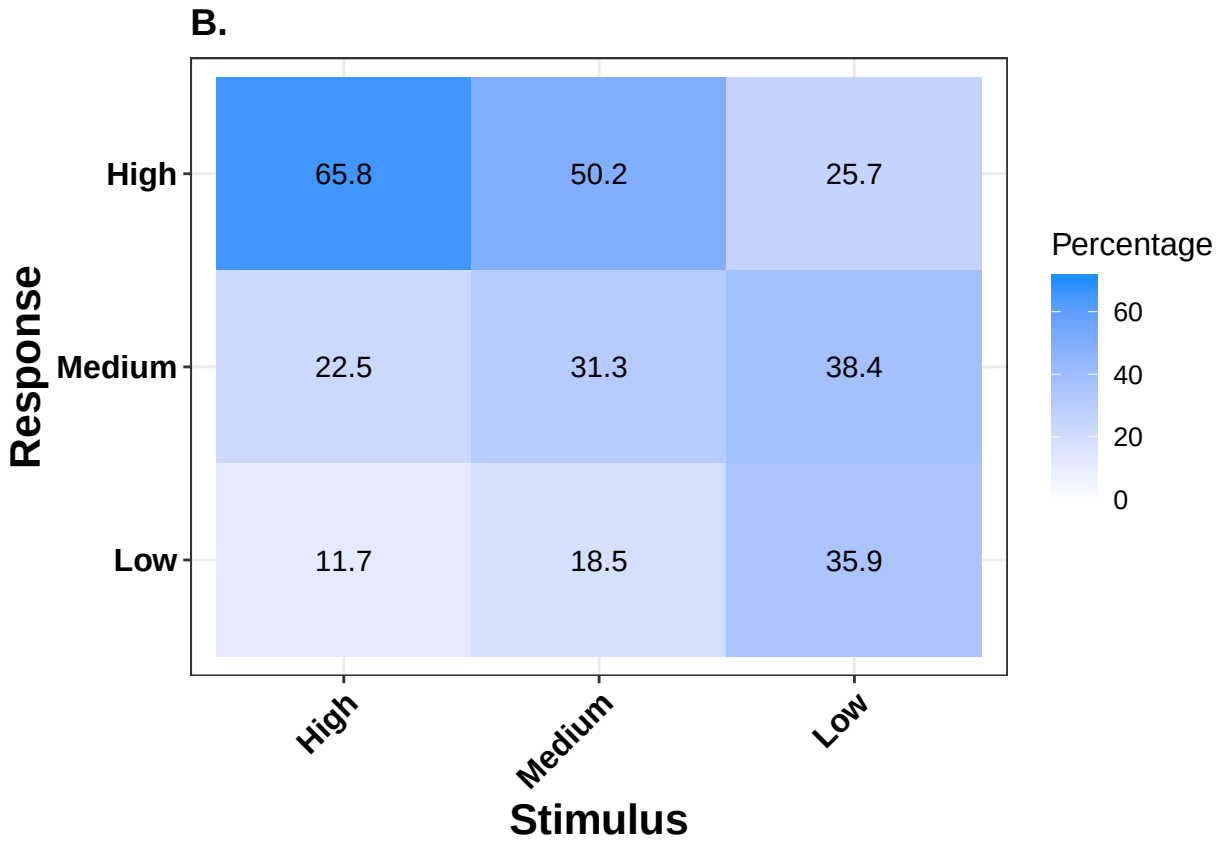
1) Confusion Matrix for recognition of behavioral context, arousal levels and valence - Experiment 1

This code displays confusion matrices for average portion of target and nontarget categories selected for each behavioral context, arousal levels and valence.

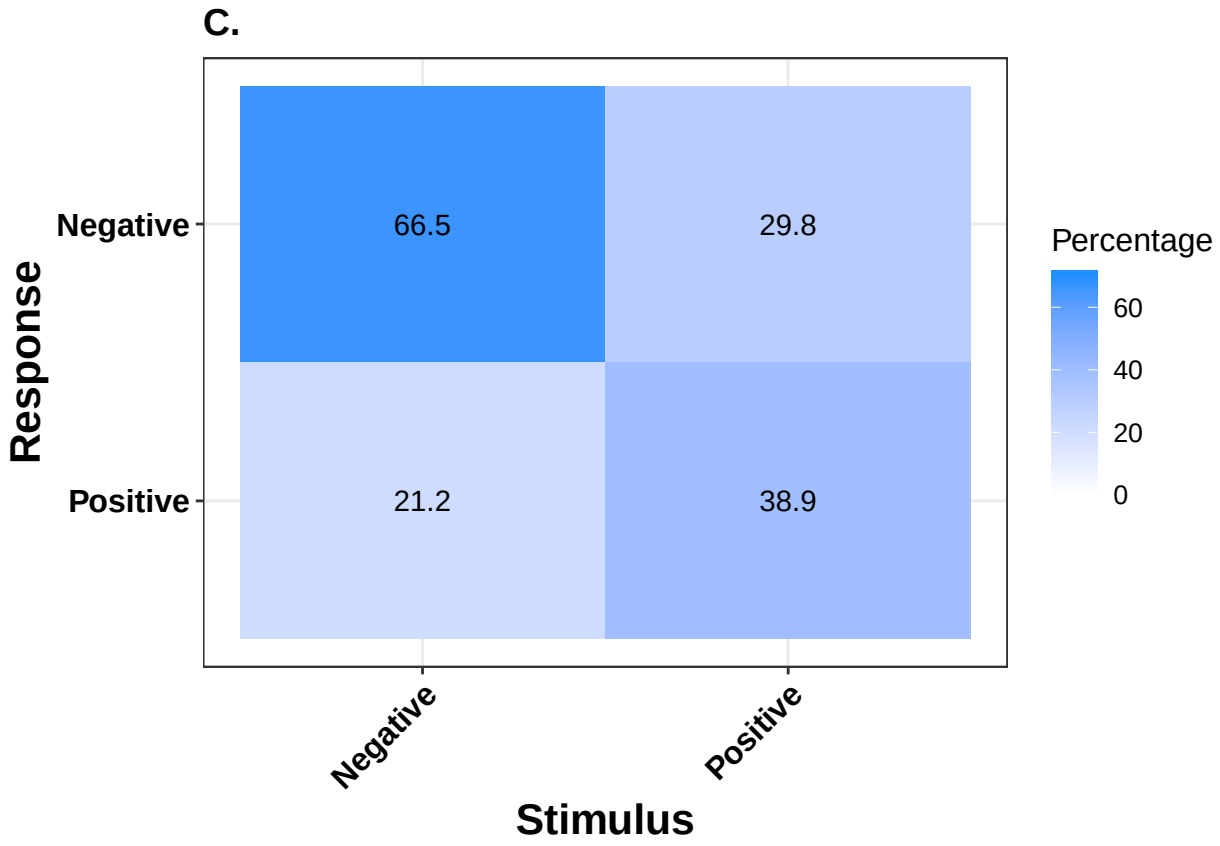
```
library(ggplot2)
library("readxl")

# Confusion matrix of behavioral contexts
confusion_matrix_context <- read_excel ("confusion_matrix_context.xlsx")
confusion_matrix_context$response<-factor(confusion_matrix_context$response, levels = c("c10","c9","c8"
confusion_matrix_context$stimulus<-factor(confusion_matrix_context$stimulus, levels = c("c1","c2","c3",
a = ggplot(confusion_matrix_context, aes(factor(stimulus), factor(response))) + geom_tile(aes(fill = Pe
a
```




```
# Confusion matrix of arousal
confusion_matrix_arousal <- read_excel ("confusion_matrix_arousal.xlsx")
confusion_matrix_arousal$response<-factor(confusion_matrix_arousal$response, levels = c("Low","Medium",
confusion_matrix_arousal$stimulus<-factor(confusion_matrix_arousal$stimulus, levels = c("High","Medium"
b = ggplot(confusion_matrix_arousal, aes(factor(stimulus), factor(response))) + geom_tile(aes(fill = Pe
b
```




```
# Confusion matrix of valence
confusion_matrix_valence <- read_excel ("confusion_matrix_valence.xlsx")
confusion_matrix_valence$response<-factor(confusion_matrix_valence$response, levels = c("Positive", "Neg
confusion_matrix_valence$stimulus<-factor(confusion_matrix_valence$stimulus, levels = c("Negative","Posi
c = ggplot(confusion_matrix_valence, aes(factor(stimulus), factor(response))) + geom_tile(aes(fill = Per
c
```



2) D prime index per behavioural context (Experiment 2)

We can illustrate d prime scores per behavioural context.

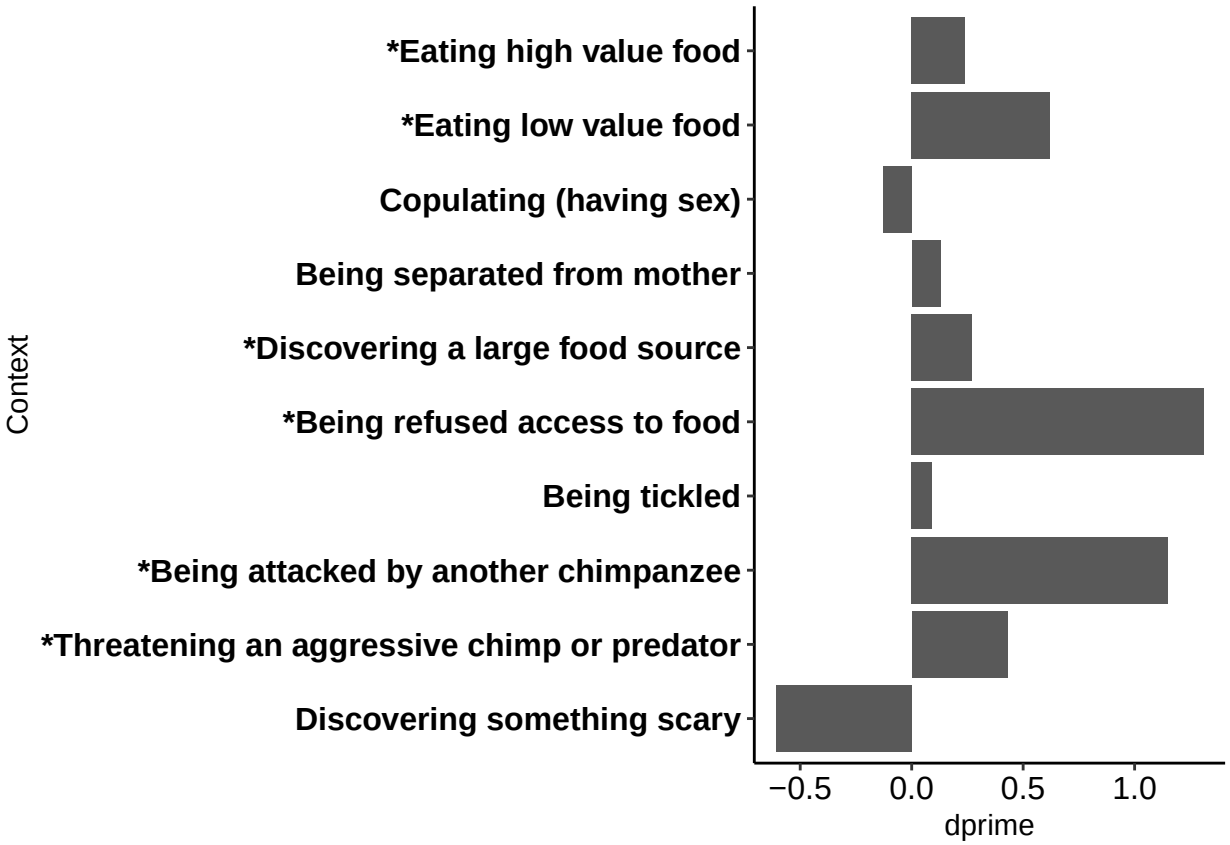
```
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library("readxl")
# D prime illustration per context
dprime <- read_excel ("dprime.xlsx")
dprime$Context<-factor(dprime$Context, levels = c("Discovering something scary","*Threatening an aggres

e<-ggplot(data=dprime, aes(x = factor(Context), y = dprime)) + geom_bar(stat="identity") + coord_flip()
e
```



3.1) Predicting recognition of context categories from acoutic features

Use GLMM and test which acoustic features predict humans' ability to recognize the behavioral context in which the chimpanzee vocalizations were produced:

```
library (lme4)
```

```
## Loading required package: Matrix
```

```
library("readxl")
datasc <- read_excel ("prediction_exp2_context.xlsx")
pvars <- c("SCoG", "Duration","F0Mean","F0Sd", "HNRMean", "HNRMax")
datasc[pvars] <- lapply(datasc[pvars], scale)

ContextRecog <- glmer(AccCont ~ (SCoG + Duration + F0Mean + F0Sd+ HNRMean + HNRMax) + (1 | PartID)  + (

print (ContextRecog)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: AccCont ~ (SCoG + Duration + F0Mean + F0Sd + HNRMean + HNRMax) +
##     (1 | PartID) + (1 | ChimpID)
##    Data: datasc
##       AIC       BIC    logLik  deviance  df.resid
## 17893.979 17963.284 -8937.989 17875.979     16318
## Random effects:
##  Groups  Name        Std.Dev.
```

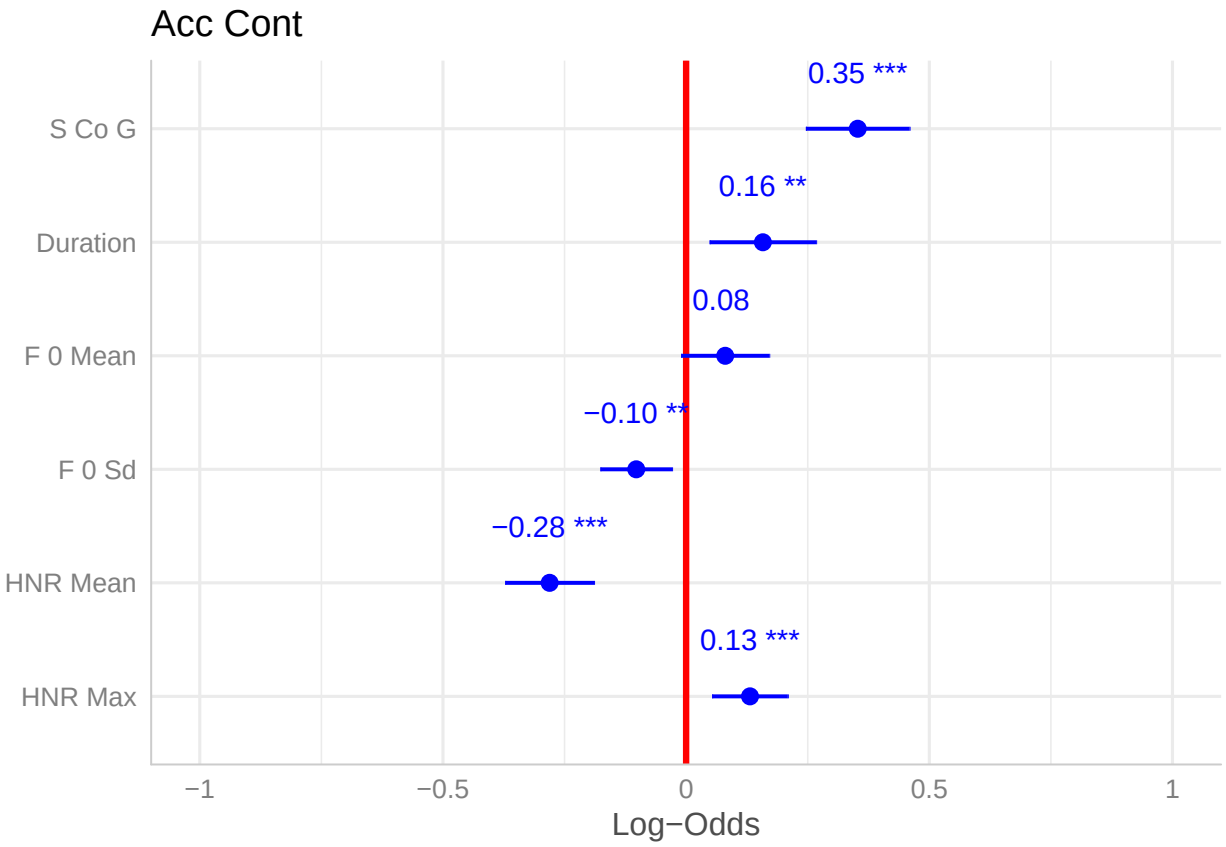
```
##  PartID  (Intercept) 1.812
##  ChimpID (Intercept) 0.522
## Number of obs: 16327, groups:  PartID, 1865; ChimpID, 51
## Fixed Effects:
## (Intercept)         SCoG     Duration       F0Mean         F0Sd
##     0.66917      0.35284      0.15780      0.08037     -0.10259
##     HNRMean       HNRMax
##    -0.28067      0.13133
```

```
summary(ContextRecog)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: AccCont ~ (SCoG + Duration + F0Mean + F0Sd + HNRMean + HNRMax) +
##     (1 | PartID) + (1 | ChimpID)
##    Data: datasc
## Control:
## glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 10000))
##
##      AIC      BIC   logLik deviance df.resid
##  17894.0  17963.3  -8938.0  17876.0    16318
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.4058 -0.5530  0.2643  0.5427  3.4882
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  PartID  (Intercept) 3.2843   1.812
##  ChimpID (Intercept) 0.2724   0.522
## Number of obs: 16327, groups:  PartID, 1865; ChimpID, 51
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.66917    0.08929   7.495 6.64e-14 ***
## SCoG         0.35284    0.05440   6.486 8.82e-11 ***
## Duration     0.15780    0.05585   2.825 0.004724 **
## F0Mean       0.08037    0.04620   1.740 0.081942 .
## F0Sd        -0.10259    0.03759  -2.729 0.006347 **
## HNRMean     -0.28067    0.04658  -6.026 1.68e-09 ***
## HNRMax       0.13133    0.03968   3.310 0.000933 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SCoG   Duratn F0Mean F0Sd   HNRMen
## SCoG      0.003
## Duration -0.078 -0.100
## F0Mean    0.032 -0.405 -0.207
## F0Sd     -0.014 -0.052 -0.253  0.060
## HNRMean  -0.061 -0.292 -0.084  0.035  0.074
## HNRMax    0.058  0.128 -0.003  0.015 -0.221 -0.572
```

We can plot the logodds of the estimates - context:

```
##
## Attaching package: 'sjlabelled'
## The following object is masked from 'package:dplyr':
##
##     as_label
```



3.2) Predicting recognition of arousal from acoutic features

Use GLMM and test which acoustic features predict humans' ability to recognize arousal in chimpanzee vocalizations:

```
library (lme4)
library("readxl")
datasc <- read_excel ("prediction_exp1_arousal.xlsx")
pvars <- c("SCoG", "Duration","F0Mean", "HNRMean")
datasc[pvars] <- lapply(datasc[pvars], scale)

ArousalRecog <- glmer(AccAro ~ (SCoG + Duration + F0Mean + HNRMean) + (1 | PartID)  + (1 | ChimpID), da

print (ArousalRecog)
```

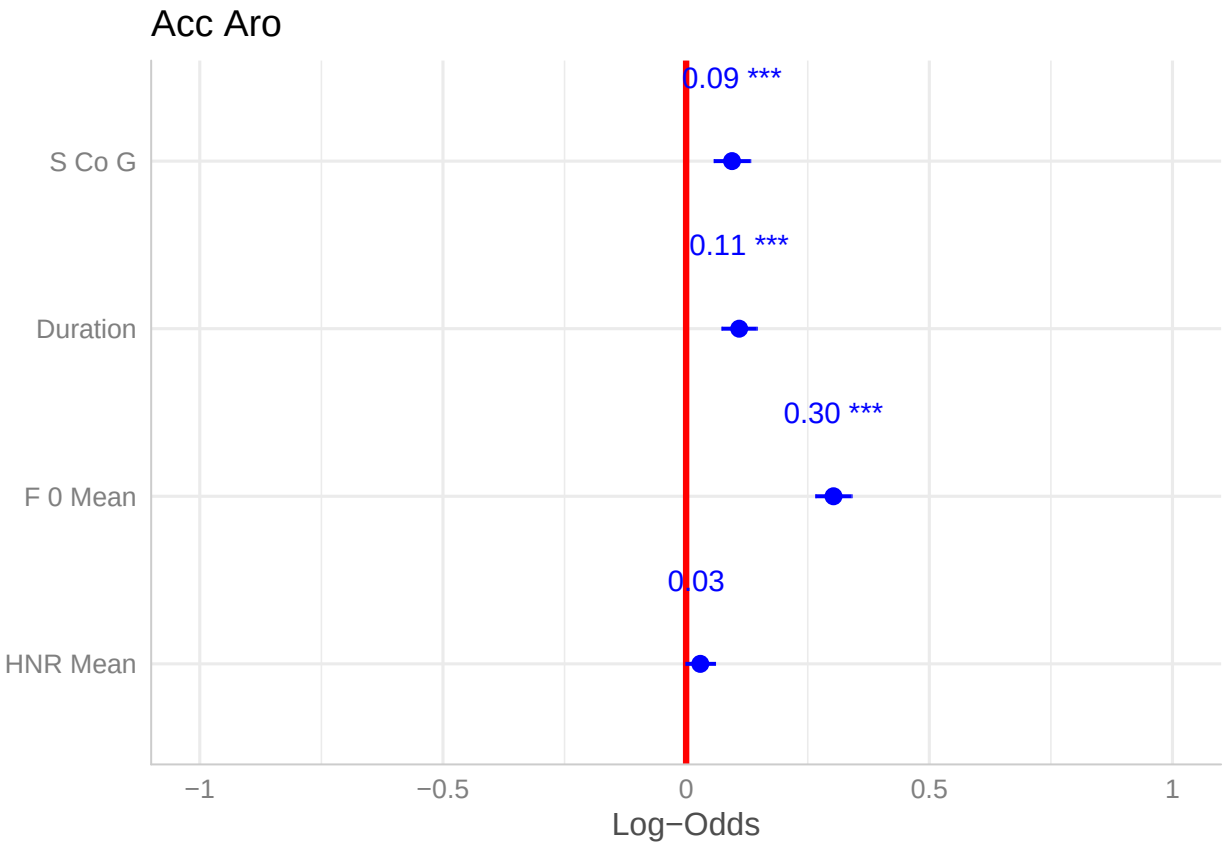
```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: AccAro ~ (SCoG + Duration + F0Mean + HNRMean) + (1 | PartID) +
##     (1 | ChimpID)
##    Data: datasc
```

```
##       AIC       BIC    logLik  deviance  df.resid
##  56340.09  56401.50 -28163.04  56326.09     47733
## Random effects:
##  Groups  Name        Std.Dev.
##  PartID  (Intercept) 0.2649
##  ChimpID (Intercept) 1.3076
## Number of obs: 47740, groups:  PartID, 310; ChimpID, 66
## Fixed Effects:
## (Intercept)         SCoG     Duration       F0Mean      HNRMean
##    -0.07938      0.09433      0.10907      0.30300      0.02915
```

```
summary(ArousalRecog)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: AccAro ~ (SCoG + Duration + F0Mean + HNRMean) + (1 | PartID) +
##     (1 | ChimpID)
##    Data: datasc
## Control:
## glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 10000))
##
##      AIC      BIC   logLik deviance df.resid
##  56340.1  56401.5 -28163.0  56326.1    47733
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3032 -0.7774 -0.2834  0.8259  7.5279
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  PartID  (Intercept) 0.07016  0.2649
##  ChimpID (Intercept) 1.70993  1.3076
## Number of obs: 47740, groups:  PartID, 310; ChimpID, 66
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.07938    0.16242  -0.489   0.6250
## SCoG         0.09433    0.01913   4.932 8.15e-07 ***
## Duration     0.10907    0.01845   5.911 3.40e-09 ***
## F0Mean       0.30300    0.01899  15.956  < 2e-16 ***
## HNRMean      0.02915    0.01545   1.887   0.0591 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SCoG   Duratn F0Mean
## SCoG      0.008
## Duration -0.013  0.065
## F0Mean   -0.016 -0.476 -0.223
## HNRMean  -0.007 -0.205 -0.322  0.166
```

We can plot the logodds of the estimates - arousal:



3.3) Predicting recognition of valence from acoutic features

Use GLMM and test which acoustic features predict humans' ability to recognize valence in chimpanzee vocalizations:

```
library (lme4)
library("readxl")
datasc <- read_excel ("prediction_exp1_valence.xlsx")
pvars <- c("SCoG", "Duration","F0Mean","F0Sd", "HNRMean", "HNRMax")
datasc[pvars] <- lapply(datasc[pvars], scale)

ValenceRecog <- glmer(AccVal ~ (SCoG + Duration + F0Mean + F0Sd+ HNRMean + HNRMax) + (1 | PartID)  + (1

print (ValenceRecog)
```

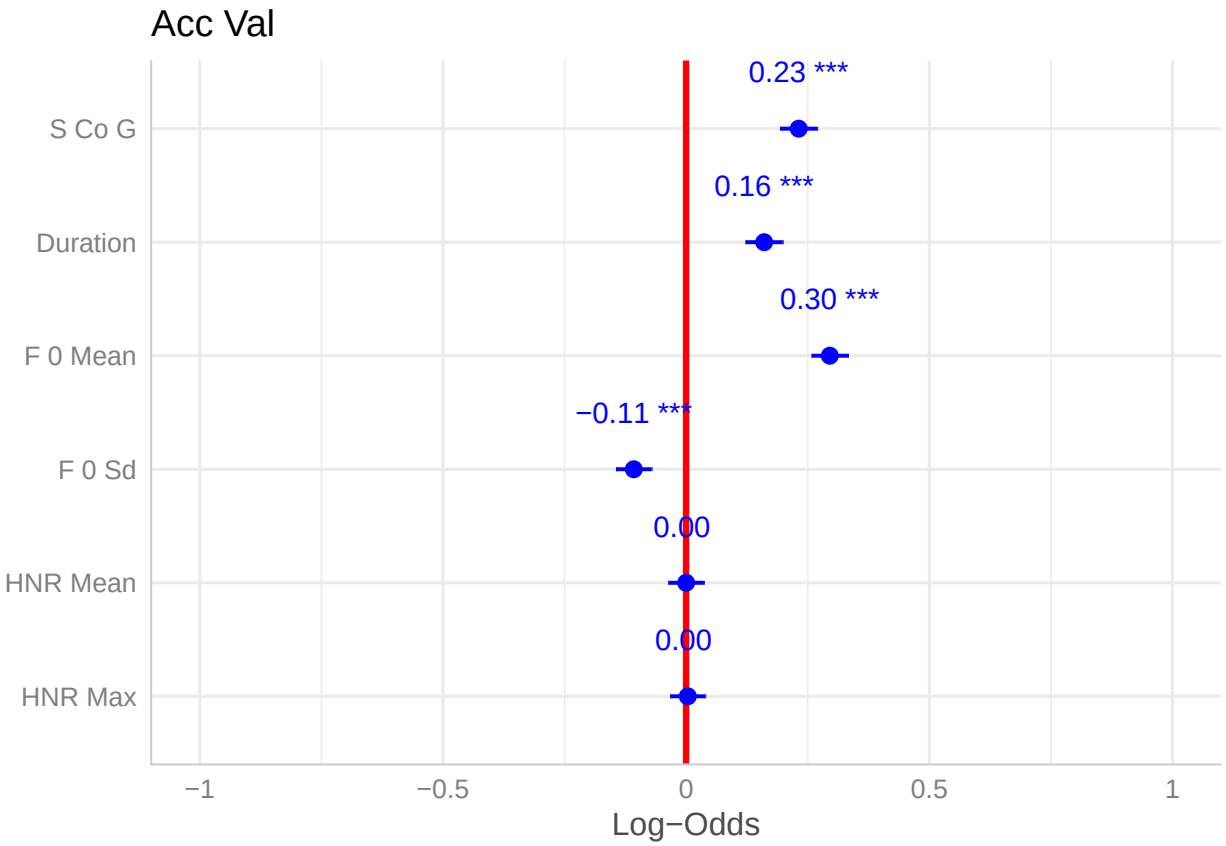
```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: AccVal ~ (SCoG + Duration + F0Mean + F0Sd + HNRMean + HNRMax) +
##     (1 | PartID) + (1 | ChimpID)
##    Data: datasc
##       AIC       BIC    logLik  deviance  df.resid
##  54158.92  54237.22 -27070.46  54140.92     44321
## Random effects:
##  Groups  Name        Std.Dev.
##  PartID  (Intercept) 0.4192
##  ChimpID (Intercept) 0.6796
```

```
## Number of obs: 44330, groups:  PartID, 310; ChimpID, 61
## Fixed Effects:
## (Intercept)         SCoG     Duration       F0Mean         F0Sd
##   2.505e-01    2.315e-01    1.605e-01    2.954e-01   -1.076e-01
##     HNRMean       HNRMax
##   2.402e-05    3.417e-03
```

```
summary(ValenceRecog)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: AccVal ~ (SCoG + Duration + F0Mean + F0Sd + HNRMean + HNRMax) +
##     (1 | PartID) + (1 | ChimpID)
##    Data: datasc
## Control:
## glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 10000))
##
##      AIC      BIC   logLik deviance df.resid
##  54158.9  54237.2 -27070.5  54140.9    44321
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5474 -0.8338  0.3756  0.8240  3.6867
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  PartID  (Intercept) 0.1757   0.4192
##  ChimpID (Intercept) 0.4619   0.6796
## Number of obs: 44330, groups:  PartID, 310; ChimpID, 61
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.505e-01  9.149e-02   2.738  0.00618 **
## SCoG         2.315e-01  1.936e-02  11.958  < 2e-16 ***
## Duration     1.605e-01  1.947e-02   8.242  < 2e-16 ***
## F0Mean       2.954e-01  1.908e-02  15.483  < 2e-16 ***
## F0Sd        -1.076e-01  1.862e-02  -5.777  7.6e-09 ***
## HNRMean      2.402e-05  1.861e-02   0.001  0.99897
## HNRMax       3.417e-03  1.836e-02   0.186  0.85234
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##          (Intr) SCoG   Duratn F0Mean F0Sd   HNRMen
## SCoG      0.020
## Duration -0.023  0.064
## F0Mean   -0.026 -0.460 -0.198
## F0Sd     -0.020 -0.045 -0.148 -0.116
## HNRMean  -0.030 -0.227 -0.205  0.124  0.105
## HNRMax    0.033  0.100 -0.123  0.018 -0.225 -0.569
```

We can plot the logodds of the estimates - valence:


Acc Val
S Co G
Duration
F 0 Mean
F 0 Sd
HNR Mean
HNR Max
0.23 ***
0.16 ***
0.30 ***
−0.11 ***
0.00
0.00
−1
−0.5
0
0.5
1
Log−Odds