Analysis of effects of bilateral cortical inactivation on sound localization using logistic regression

Version History:

- Created: 2021-09-11 Stephen Town
- Ported to notebook: 2022-05-16

Combine data across ferrets and run a generalized linear mixed model to determine the effect of cooling on probability of making a 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))
```

---

1. Loading and preprocessing

Load data from individual subjects and combine.

```
data_path = '/home/stephen/Github/Vowel_Discrimination_In_Noise/Results/Localization/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)

magnum$fNum <- 1311
robin$fNum  <- 1509

df <- rbind(magnum, robin) %>%
  mutate(
    fNum = as.factor(fNum),         # Treat F number as
    HoldTime = HoldTime / 1000      # Convert ms to seconds
    )

df <- subset(df, select=-c(originalFile, SessionDate))      # Remove information about data origin (un
df <- subset(df, select=-c(Trial, StartTime, ComputerTime))  # Remove information about trial order (un

rm(magnum, robin)
head(df)
```

```
## # A tibble: 6 x 21
##   CorrectionTrial CenterReward SpkrIdx SpkrPos duration HoldTime Atten Response
##             <dbl>        <dbl>   <dbl>   <dbl>    <dbl>    <dbl> <dbl>    <dbl>
```

```
## 1              0          0       7    -60      250   0.901  0           7
## 2              0          0       4     30      250   0.601  4.5         4
## 3              0          0       6    -30      250   1.00   0           6
## 4              0          0       7    -60      250   0.651  9           7
## 5              0          0       2     90      250   1.00   9           3
## 6              0          0       3     60      250   0.801  9           3
## # ... with 13 more variables: RespTime <dbl>, Correct <dbl>, Modality <chr>,
## #   Condition <chr>, LeftTemperature <dbl>, RightTemperature <dbl>, fNum <fct>,
## #   ResponseAngle <dbl>, Spkr_Rad <dbl>, spkr_hemifield <chr>,
## #   resp_hemifield <chr>, LatCorrect <dbl>, Error <dbl>
```

Rescale speaker angle and response error from degrees to radians, and remove older ways of encoding sound and response location (in degrees or as index)

```r
df <- df %>%
  mutate(
    SpeakerRad = SpkrPos / 180 * pi,
    ErrorRad = Error / 180 *pi
    ) %>%
  select(-SpkrIdx, -SpkrPos, -Response, -ResponseAngle, Error)


head(df)
```

```
## # A tibble: 6 x 19
##   CorrectionTrial CenterReward duration HoldTime Atten RespTime Correct Modality
##             <dbl>        <dbl>    <dbl>    <dbl> <dbl>    <dbl>   <dbl> <chr>
## 1              0            0      250    0.901  0       1068.       1 Aud
## 2              0            0      250    0.601  4.5     1085.       1 Aud
## 3              0            0      250    1.00   0       1102.       1 Aud
## 4              0            0      250    0.651  9       1123.       1 Aud
## 5              0            0      250    1.00   9       1223.       0 Aud
## 6              0            0      250    0.801  9       1248.       1 Aud
## # ... with 11 more variables: Condition <chr>, LeftTemperature <dbl>,
## #   RightTemperature <dbl>, fNum <fct>, Spkr_Rad <dbl>, spkr_hemifield <chr>,
## #   resp_hemifield <chr>, LatCorrect <dbl>, Error <dbl>, SpeakerRad <dbl>,
## #   ErrorRad <dbl>
```

Aggregate counts across unique trials conditions... converting binary responses to count data can make it easier to perform checks and further predictions on. Note that the coefficients for the model should be the same as the model fitted to binary data above.

```r
counts = df %>%
  group_by(fNum, Condition, Atten, SpeakerRad, spkr_hemifield, CenterReward) %>%
  summarise(
    nCorrect = sum(Correct),
    nTotal = n(),
    .groups = 'keep'
    ) %>%
  mutate(
    dist_to_midline = abs(SpeakerRad)
  )


bc_counts <- counts[counts$Condition %in% c("Bilateral","Control"),]
```

3. Fitting models to test effect of Bilateral cooling on accuracy

3.1. Model fitting

Start by doing the logistic regression on performance on full localization. Here, we're interested in the effect of condition (bilateral cooling or control) on odds of making a correct response. For completeness, we'll include covariates (signal attenuation, speaker location, center reward) where possible, to try and account for as much variance in the data as possible. Note that there are other variables (Hold Time and sound duration) that we might ideally like to include as further covariates, but these are confounded with subject and so may be the cause of 'isSingular warnings'.

It's also important to consider that non-linear relationships may exist between speaker angle and performance. To accommodate this, we will look at a couple of ways of including sound location as part of the model.

Model 1: Ignore speaker completely

```
BCC_1 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition + Atten + CenterReward + (1|fNum),
  family=binomial,
  data=bc_counts)

summary(BCC_1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition + Atten +
##     CenterReward + (1 | fNum)
##    Data: bc_counts
##
##      AIC      BIC   logLik deviance df.resid
##    848.8    863.3   -419.4    838.8      129
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.1226 -0.9735 -0.1285  1.2517  5.7179
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.02171  0.1473
## Number of obs: 134, groups:  fNum, 2
##
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -0.276803   0.152174  -1.819 0.068913 .
## ConditionControl  0.585653   0.108713   5.387 7.16e-08 ***
## Atten            -0.035538   0.009659  -3.679 0.000234 ***
## CenterReward     -0.493677   0.147620  -3.344 0.000825 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) CndtnC Atten
## CondtnCntrl -0.624
## Atten      -0.266 -0.031
## CenterRewrd -0.063 -0.010  0.015
```

Model 2: Consider model as a random effect

```
BCC_2 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition + Atten + CenterReward + (1|fNum) + (1|SpeakerRad)
  family=binomial,
  data=bc_counts)

summary(BCC_2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition + Atten +
##     CenterReward + (1 | fNum) + (1 | SpeakerRad)
##    Data: bc_counts
##
##      AIC      BIC   logLik deviance df.resid
##    752.3    769.7   -370.2    740.3      128
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.3826 -0.8010 -0.1721  1.1488  4.1365
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  SpeakerRad (Intercept) 0.23171  0.4814
##  fNum       (Intercept) 0.07048  0.2655
## Number of obs: 134, groups:  SpeakerRad, 7; fNum, 2
##
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -0.385321   0.285079  -1.352 0.176494
## ConditionControl  0.624949   0.110901   5.635 1.75e-08 ***
## Atten            -0.036171   0.009861  -3.668 0.000244 ***
## CenterReward     -0.511796   0.149972  -3.413 0.000643 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) CndtnC Atten
## CondtnCntrl -0.341
## Atten       -0.144 -0.031
## CenterRewrd -0.033 -0.013  0.016
```

Model 3: Consider the distance from the midline and hemifield of the speaker, without interaction with cooling

```
BCC_3 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition + dist_to_midline*spkr_hemifield + Atten + CenterR
  family=binomial,
  data=bc_counts)

summary(BCC_3)
```

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

4

```
##  Family: binomial  ( logit )
## Formula:
## cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition + dist_to_midline *
##      spkr_hemifield + Atten + CenterReward + (1 | fNum)
##    Data: bc_counts
##
##      AIC      BIC   logLik deviance df.resid
##    732.4    754.8   -358.2    716.4      114
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6029 -1.0090 -0.0705  1.2347  5.0179
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.05343  0.2311
## Number of obs: 122, groups:  fNum, 2
##
## Fixed effects:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    -0.54766    0.24272  -2.256  0.02405 *
## ConditionControl                0.62937    0.11451   5.496 3.89e-08 ***
## dist_to_midline                 0.31316    0.12605   2.484  0.01298 *
## spkr_hemifieldR                -0.10638    0.19918  -0.534  0.59329
## Atten                          -0.03074    0.01017  -3.023  0.00251 **
## CenterReward                   -0.48172    0.15160  -3.178  0.00149 **
## dist_to_midline:spkr_hemifieldR  0.01987    0.17635   0.113  0.91031
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) CndtnC dst_t_ spkr_R Atten  CntrRw
## CondtnCntrl -0.415
## dist_t_mdln -0.544  0.008
## spkr_hmfldR -0.407 -0.011  0.668
## Atten       -0.149 -0.032 -0.037 -0.061
## CenterRewrd -0.024 -0.016 -0.032 -0.010  0.018
## dst_t_md:_R  0.378  0.012 -0.715 -0.929  0.049  0.018
```

Make the model slightly more complex by considering the distance to the midline and hemifield, with an interaction

```
BCC_4 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition*dist_to_midline + dist_to_midline*spkr_hemifield +
  family=binomial,
  data=bc_counts)

summary(BCC_4)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula:
## cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition * dist_to_midline +
##     dist_to_midline * spkr_hemifield + Atten + CenterReward +      (1 | fNum)
```

```
##    Data: bc_counts
##
##     AIC     BIC  logLik deviance df.resid
##    722.6   747.8  -352.3   704.6      113
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.3739 -1.0060  0.0645  1.3835  4.7738
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.05379  0.2319
## Number of obs: 122, groups:  fNum, 2
##
## Fixed effects:
##                                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)                     -1.43421    0.36239  -3.958 7.57e-05 ***
## ConditionControl                 1.63770    0.32336   5.065 4.09e-07 ***
## dist_to_midline                  1.13315    0.27485   4.123 3.74e-05 ***
## spkr_hemifieldR                 -0.11633    0.20046  -0.580 0.561701
## Atten                           -0.03047    0.01019  -2.991 0.002781 **
## CenterReward                    -0.48613    0.15164  -3.206 0.001346 **
## ConditionControl:dist_to_midline -0.93514   0.27666  -3.380 0.000725 ***
## dist_to_midline:spkr_hemifieldR  0.02500    0.17687   0.141 0.887617
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) CndtnC dst_t_ spkr_R Atten  CntrRw CnC:__
## CondtnCntrl -0.791
## dist_t_mdln -0.823  0.831
## spkr_hmfldR -0.263 -0.019  0.294
## Atten       -0.103 -0.008 -0.013 -0.061
## CenterRewrd -0.009 -0.015 -0.024 -0.011  0.018
## CndtnCnt:__  0.736 -0.932 -0.888  0.015 -0.004  0.010
## dst_t_md:_R  0.248  0.013 -0.321 -0.929  0.049  0.019 -0.009
```

Finally, consider the three way interaction between speaker hemifield, distance to midline and cooling condition

```
BCC_5 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition*dist_to_midline*spkr_hemifield + Atten + CenterRewa
  family=binomial,
  data=bc_counts)

summary(BCC_5)
```

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

```
##     693.1    723.9    -335.5     671.1        111
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9755 -0.9987 -0.2170  1.3051  4.8403
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.05354  0.2314
## Number of obs: 122, groups:  fNum, 2
##
## Fixed effects:
##                                                 Estimate Std. Error z value
## (Intercept)                                     -3.09152    0.54027  -5.722
## ConditionControl                                 3.48173    0.53608   6.495
## dist_to_midline                                  2.73690    0.43916   6.232
## spkr_hemifieldR                                  2.88770    0.65505   4.408
## Atten                                           -0.03009    0.01024  -2.938
## CenterReward                                    -0.51531    0.15248  -3.380
## ConditionControl:dist_to_midline                -2.72537    0.45908  -5.937
## ConditionControl:spkr_hemifieldR                -3.36667    0.68833  -4.891
## dist_to_midline:spkr_hemifieldR                 -2.89281    0.56437  -5.126
## ConditionControl:dist_to_midline:spkr_hemifieldR 3.28493    0.59479   5.523
##                                                 Pr(>|z|)
## (Intercept)                                     1.05e-08 ***
## ConditionControl                                8.32e-11 ***
## dist_to_midline                                 4.60e-10 ***
## spkr_hemifieldR                                 1.04e-05 ***
## Atten                                           0.003307 **
## CenterReward                                    0.000726 ***
## ConditionControl:dist_to_midline                2.91e-09 ***
## ConditionControl:spkr_hemifieldR                1.00e-06 ***
## dist_to_midline:spkr_hemifieldR                 2.96e-07 ***
## ConditionControl:dist_to_midline:spkr_hemifieldR 3.34e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) CndtnC dst_t_ spkr_R Atten  CntrRw CnC:__ CnC:_R d__:_R
## CondtnCntrl -0.910
## dist_t_mdln -0.895  0.903
## spkr_hmfldR -0.743  0.751  0.739
## Atten       -0.062 -0.013 -0.010 -0.030
## CenterRewrd  0.012 -0.029 -0.040 -0.026  0.018
## CndtnCnt:__  0.856 -0.941 -0.956 -0.706 -0.001  0.033
## CndtnCnt:_R  0.708 -0.778 -0.703 -0.951  0.012  0.025  0.733
## dst_t_md:_R  0.695 -0.703 -0.778 -0.937  0.018  0.036  0.744  0.891
## CndtC:__:_R -0.661  0.726  0.738  0.888 -0.003 -0.034 -0.772 -0.936 -0.949
```

Model comparison

```
anova(BCC_3, BCC_4, BCC_5)
```

```
## Data: bc_counts
## Models:
```

```
## BCC_3: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition + dist_to_midline * spkr_hemifield + Atten
## BCC_4: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition * dist_to_midline + dist_to_midline * spkr_
## BCC_5: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition * dist_to_midline * spkr_hemifield + Atten
##       npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## BCC_3    8 732.39 754.82 -358.19   716.39
## BCC_4    9 722.61 747.85 -352.31   704.61 11.778  1  0.0005992 ***
## BCC_5   11 693.06 723.90 -335.53   671.06 33.555  2  5.171e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.2. Model checking

Checking logistic regression models is difficult because many of the conventional indicators of poor model quality used in linear regression (heteroscedacity, Gaussian distribution of residuals) don't apply. Here, we'll use the DHARMa package to calculate randomized quantile residuals using a simulation approach. This hypothesis is that the cumulative distribution of residuals should give average values close to 0.5 and within the bounds of 0.25 to 0.75. We can inspect how residuals differ between conditions to get an idea of where problems might be occuring ()

```
simulationOutput <- simulateResiduals(fittedModel = BCC_2, n=500, plot = T)
```
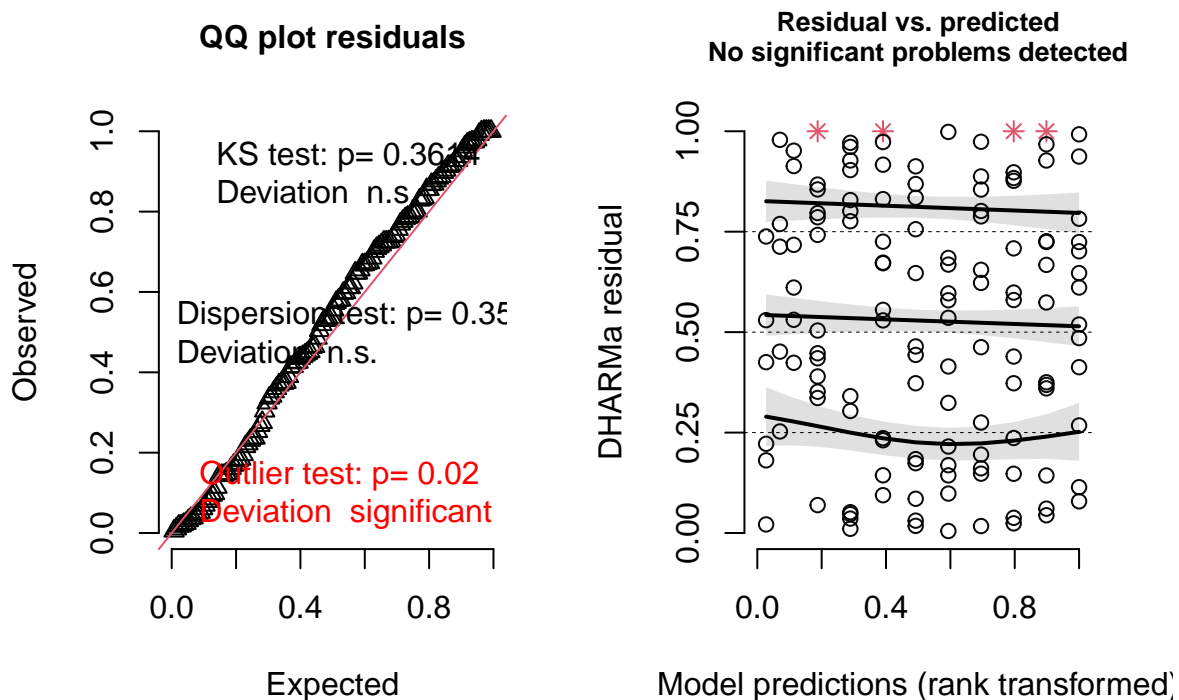

3.3. Model Prediction

Ok, let's try instead using **model prediction** to see if the process generated by the model matches what we see in the data. We start by adding a column to the input table that consists of the probability of making a

correct response according to the model, from which we then estimate the number of correct trials for each type of experimental condition.

```r
# Predict probability of making correct response
bc_counts$fit <- predict(BCC_2, bc_counts, type="response")

# Predict the number of correct trials, using original trial counts
bc_counts <- bc_counts %>%
  mutate(
    predicted_correct = fit * nTotal)

# Summarize data center reward and attenuation
bc_agg = bc_counts %>%
  group_by(fNum, Condition, SpeakerRad) %>%
  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
  )
```

We could also make a prediction about the behavior across sound locations as was shown in the initial visualization. Let's therefore Aggregate across variables that aren't of interest and plot the observed behavior vs. the predicted behavior.

```r
ggplot(                                         # Plot facet grid for predicted and observed performance
  bc_agg %>%
    mutate(Spkr_deg = SpeakerRad / pi * 180),
  aes(x = Spkr_deg, color=Condition)
  ) +
geom_line(
  aes(y = fit_p)
  ) +
geom_point(
  shape = 19,
  aes(
    y = pCorrect,
    size = nTotal,
    alpha = 0.25
    )
  ) +
scale_size(range = c(0, 4), trans="log10") +
scale_color_manual(
  labels = c("Cooled", "Control"),
  values = c("#37a9a4", "#3d3d3d")
) +
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")
  ) +
scale_x_continuous(name="Speaker Angle (°)", limits=c(-100, 100), breaks=seq(-90,90,30)) +
scale_y_continuous(name="% Correct", limits=c(0, 100)) +
facet_wrap(. ~fNum)
```
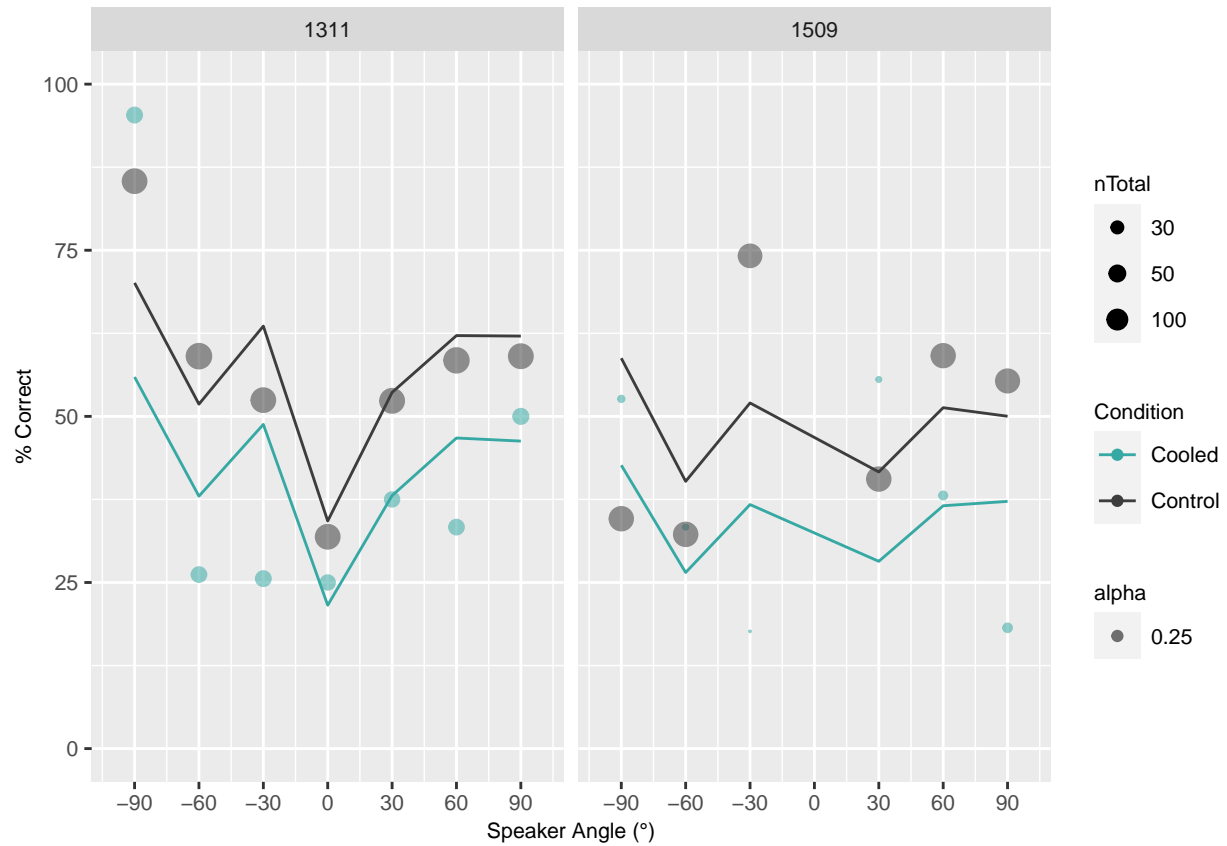


```
ggsave(filename = "Localization_predictions_BCC2b.png", width = 12, height = 6, units = "cm", dpi=300)
```