Analysis of effects of unilateral 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))
suppressPackageStartupMessages( library(sjPlot))
```

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 (unless neede
d for debugging)
df <- subset(df, select=-c(Trial, StartTime, ComputerTime))  # Remove information about trial order (unless neede
d for debugging)

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

CorrectionTrial <dbl>	CenterReward <dbl>	SpkrIdx <dbl>	SpkrPos <dbl>	duration <dbl>	HoldTime <dbl>	Atten <dbl>	Response <dbl>	RespTime <dbl>	Correct <dbl>
0	0	7	-60	250	0.901	0.0	7	1068.154	1
0	0	4	30	250	0.601	4.5	4	1085.177	1
0	0	6	-30	250	1.001	0.0	6	1101.935	1
0	0	7	-60	250	0.651	9.0	7	1122.504	1
0	0	2	90	250	1.001	9.0	3	1223.321	0
0	0	3	60	250	0.801	9.0	3	1247.787	1

6 rows | 1-10 of 21 columns

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)

```
df <- df %>%
  mutate(
    SpeakerRad = SpkrPos / 180 * pi,
    ErrorRad = Error / 180 *pi,
    dist_to_midline = abs(SpkrPos) / 180 * pi
    )

head(df)
```

CorrectionTrial <dbl>	CenterReward <dbl>	SpkrIdx <dbl>	SpkrPos <dbl>	duration <dbl>	HoldTime <dbl>	Atten <dbl>	Response <dbl>	RespTime <dbl>	Correct <dbl>
0	0	7	-60	250	0.901	0.0	7	1068.154	1
0	0	4	30	250	0.601	4.5	4	1085.177	1
0	0	6	-30	250	1.001	0.0	6	1101.935	1
0	0	7	-60	250	0.651	9.0	7	1122.504	1
0	0	2	90	250	1.001	9.0	3	1223.321	0
0	0	3	60	250	0.801	9.0	3	1247.787	1

6 rows | 1-10 of 24 columns

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.

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


# Separate out trials with unilateral cooling
lr_counts <- counts[counts$Condition %in% c("Left","Right"),]

# Include performance on control trials as a covariate
control_counts = subset(counts, Condition %in% c("Control"))
control_counts = control_counts %>%
  mutate( control_pCorrect = nCorrect / nTotal) %>%
  select( control_pCorrect)
```

```
## Adding missing grouping variables: `fNum`, `Condition`, `Atten`,
## `spkr_hemifield`, `dist_to_midline`, `CenterReward`
```

```
lr_extended = merge(
  x = lr_counts,
  y = control_counts,
  by = c("fNum", "Atten", "spkr_hemifield", "dist_to_midline", "CenterReward"),
  all.x = TRUE,
  all.y = FALSE)

lr_extended = lr_extended %>%
  select( !Condition.y) %>%
  rename(Condition = Condition.x)

head(lr_extended)
```

	fN… <fct>	Atten <dbl>	spkr_hemifield <chr>	dist_to_midline <dbl>	CenterReward <dbl>	Condition <chr>	nCorrect <dbl>	nTotal <int>	control_pCorrect <dbl>
1	1311	0	L	0.5235988	0	Left	11	19	0.5609756
2	1311	0	L	0.5235988	0	Right	7	22	0.5609756
3	1311	0	L	0.5235988	1	Left	1	1	0.8000000
4	1311	0	L	0.5235988	1	Right	0	1	0.8000000

	fN…	Atten	spkr_hemifield	dist_to_midline	CenterReward	Condition	nCorrect	nTotal	control_pCorrect
	<fct>	<dbl>	<chr>	<dbl>	<dbl>	<chr>	<dbl>	<int>	<dbl>
5	1311	0	L	1.0471976	0	Right	12	24	0.5270270
6	1311	0	L	1.0471976	0	Left	15	25	0.5270270

6 rows

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 interaction between cooled hemisphere and speaker hemifield.

```
LR_1 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition*spkr_hemifield + Atten + CenterReward + (1|fNum),
  family=binomial,
  data=lr_extended)

summary(LR_1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition * spkr_hemifield +
##     Atten + CenterReward + (1 | fNum)
##    Data: lr_extended
##
##      AIC      BIC   logLik deviance df.resid
##    514.7    533.7   -250.3    500.7      105
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5085 -0.9746 -0.5714  0.9477  5.1296
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.1689   0.411
## Number of obs: 112, groups:  fNum, 2
##
## Fixed effects:
##                               Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    0.20867    0.32910   0.634   0.5260
## ConditionRight                -0.43611    0.16886  -2.583   0.0098 **
## spkr_hemifieldR               -0.99829    0.18669  -5.347 8.93e-08 ***
## Atten                         -0.02407    0.01586  -1.518   0.1290
## CenterReward                  -0.23597    0.26933  -0.876   0.3810
## ConditionRight:spkr_hemifieldR  1.14522    0.23942   4.783 1.72e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) CndtnR spkr_R Atten  CntrRw
## ConditnRght -0.308
## spkr_hmfldR -0.274  0.529
## Atten       -0.220 -0.032  0.028
## CenterRewrd -0.041 -0.050 -0.031  0.017
## CndtnRgh:_R  0.199 -0.691 -0.780  0.027  0.050
```

Model 2: It might be helpful to include the performance in control conditions

```
LR_2 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition*spkr_hemifield + Atten + CenterReward + control_pCorrect + (
1|fNum),
  family=binomial,
  data=lr_extended)

summary(LR_2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition * spkr_hemifield +
##     Atten + CenterReward + control_pCorrect + (1 | fNum)
##    Data: lr_extended
##
##      AIC      BIC   logLik deviance df.resid
##    458.4    480.2   -221.2    442.4      104
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.5604 -1.0643 -0.3020  0.9123  3.0078
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.0718   0.268
## Number of obs: 112, groups:  fNum, 2
##
## Fixed effects:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    -1.55079    0.34423  -4.505 6.64e-06 ***
## ConditionRight                 -0.45411    0.17566  -2.585  0.00973 **
## spkr_hemifieldR                -0.87410    0.19108  -4.574 4.77e-06 ***
## Atten                          -0.00260    0.01637  -0.159  0.87382
## CenterReward                   -0.07623    0.30328  -0.251  0.80155
## control_pCorrect                2.89804    0.39578   7.322 2.44e-13 ***
## ConditionRight:spkr_hemifieldR  1.15119    0.24501   4.699 2.62e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) CndtnR spkr_R Atten  CntrRw cntr_C
## ConditnRght -0.279
## spkr_hmfldR -0.323  0.535
## Atten       -0.328 -0.033  0.028
## CenterRewrd -0.061 -0.046 -0.037  0.019
## cntrl_pCrrc -0.696 -0.038  0.071  0.176  0.039
## CndtnRgh:_R  0.181 -0.704 -0.775  0.029  0.056  0.029
```

Model 3: Add in the distance from the midline

```
LR_3 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition*spkr_hemifield + dist_to_midline*spkr_hemifield + Atten + Ce
nterReward  + control_pCorrect+ (1|fNum),
  family=binomial,
  data=lr_extended)

summary(LR_3)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition * spkr_hemifield +
##     dist_to_midline * spkr_hemifield + Atten + CenterReward +
##     control_pCorrect + (1 | fNum)
##    Data: lr_extended
##
##      AIC      BIC   logLik deviance df.resid
##    421.0    448.2   -200.5    401.0      102
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.1486 -0.9724 -0.1664  0.7764  3.5159
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.08509  0.2917
## Number of obs: 112, groups:  fNum, 2
##
## Fixed effects:
##                                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    -2.574175   0.397101  -6.482 9.03e-11 ***
## ConditionRight                 -0.494614   0.183461  -2.696  0.00702 **
## spkr_hemifieldR                 0.767658   0.348151   2.205  0.02746 *
## dist_to_midline                 1.306133   0.218277   5.984 2.18e-09 ***
## Atten                          -0.008965   0.016785  -0.534  0.59326
## CenterReward                   -0.128528   0.293659  -0.438  0.66162
## control_pCorrect                2.493706   0.414370   6.018 1.77e-09 ***
## ConditionRight:spkr_hemifieldR  1.196610   0.250321   4.780 1.75e-06 ***
## spkr_hemifieldR:dist_to_midline -1.668249   0.291428  -5.724 1.04e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) CndtnR spkr_R dst_t_ Atten  CntrRw cntr_C CnR:_R
## ConditnRght -0.228
## spkr_hmfldR -0.461  0.265
## dist_t_mdln -0.442 -0.069  0.587
## Atten       -0.262 -0.027 -0.018 -0.064
## CenterRewrd -0.053 -0.057 -0.045 -0.036  0.026
## cntrl_pCrrc -0.540 -0.019  0.022 -0.134  0.190  0.070
## CndtnRgh:_R  0.151 -0.720 -0.395  0.057  0.027  0.063  0.012
## spkr_hmR:__  0.383  0.052 -0.828 -0.737  0.037  0.030  0.005 -0.051
```

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

```
LR_4 = glmer(
  cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition*spkr_hemifield*dist_to_midline + Atten + CenterReward + cont
rol_pCorrect + (1|fNum),
  family=binomial,
  data=lr_extended)

summary(LR_4)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: binomial  ( logit )
## Formula: cbind(nCorrect, nTotal - nCorrect) ~ 1 + Condition * spkr_hemifield *
##     dist_to_midline + Atten + CenterReward + control_pCorrect +      (1 | fNum)
##    Data: lr_extended
##
##      AIC      BIC   logLik deviance df.resid
##    399.4    432.0   -187.7    375.4      100
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.4490 -0.8591 -0.2736  0.8090  3.8872
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  fNum   (Intercept) 0.09371  0.3061
## Number of obs: 112, groups:  fNum, 2
##
## Fixed effects:
##                                             Estimate Std. Error z value
## (Intercept)                                -3.21463    0.49553  -6.487
## ConditionRight                              0.61722    0.49198   1.255
## spkr_hemifieldR                             2.63632    0.53064   4.968
## dist_to_midline                             2.06012    0.39007   5.281
## Atten                                      -0.01081    0.01698  -0.637
## CenterReward                               -0.08180    0.29585  -0.276
## control_pCorrect                            2.36190    0.41907   5.636
## ConditionRight:spkr_hemifieldR             -1.76877    0.66938  -2.642
## ConditionRight:dist_to_midline             -1.13814    0.47045  -2.419
## spkr_hemifieldR:dist_to_midline            -3.58141    0.51629  -6.937
## ConditionRight:spkr_hemifieldR:dist_to_midline  2.96526    0.63170   4.694
##                                            Pr(>|z|)
## (Intercept)                                8.75e-11 ***
## ConditionRight                              0.20964
## spkr_hemifieldR                            6.76e-07 ***
## dist_to_midline                            1.28e-07 ***
## Atten                                       0.52431
## CenterReward                                0.78218
## control_pCorrect                           1.74e-08 ***
## ConditionRight:spkr_hemifieldR              0.00823 **
## ConditionRight:dist_to_midline              0.01555 *
## spkr_hemifieldR:dist_to_midline            4.01e-12 ***
## ConditionRight:spkr_hemifieldR:dist_to_midline 2.68e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) CndtnR spkr_R dst_t_ Atten  CntrRw cntr_C CnR:_R CnR:__
## ConditnRght -0.606
## spkr_hmfldR -0.568  0.589
## dist_t_mdln -0.680  0.745  0.681
## Atten       -0.231  0.025 -0.018 -0.006
## CenterRewrd -0.073  0.021  0.006  0.018  0.029
## cntrl_pCrrc -0.382 -0.093 -0.036 -0.138  0.183  0.071
## CndtnRgh:_R  0.435 -0.732 -0.794 -0.545  0.017 -0.017  0.070
## CndtnRgh:__  0.584 -0.925 -0.562 -0.825 -0.042 -0.045  0.077  0.677
## spkr_hmR:__  0.531 -0.557 -0.924 -0.748  0.023 -0.022  0.050  0.735  0.618
## CndtR:_R:__ -0.431  0.689  0.756  0.615 -0.003  0.044 -0.060 -0.924 -0.744
##             s_R:__
## ConditnRght
## spkr_hmfldR
## dist_t_mdln
## Atten
## CenterRewrd
## cntrl_pCrrc
## CndtnRgh:_R
## CndtnRgh:__
## spkr_hmR:__
## CndtR:_R:__ -0.819
```

Model comparison

```
anova(LR_1, LR_2, LR_3, LR_4)
```

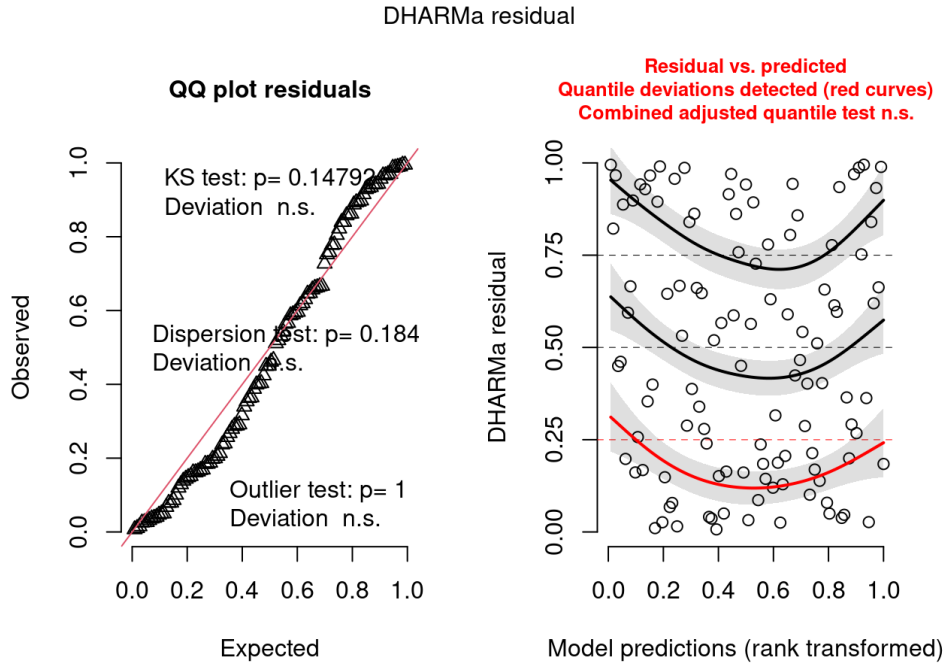
	npar <dbl>	AIC <dbl>	BIC <dbl>	logLik <dbl>	deviance <dbl>	Chisq <dbl>	Df <dbl>	Pr(>Chisq) <dbl>
LR_1	7	514.6503	533.6798	-250.3251	500.6503	*NA*	*NA*	*NA*
LR_2	8	458.4117	480.1597	-221.2058	442.4117	58.23860	1	2.321786e-14
LR_3	10	420.9771	448.1620	-200.4885	400.9771	41.43461	2	1.005979e-09
LR_4	12	399.3708	431.9928	-187.6854	375.3708	25.60629	2	2.752110e-06

4 rows

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 = LR_4, n=500, plot = T)
```




```
tab_model(LR_4)
```

	cbind(nCorrect, nTotal - nCorrect)		
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	0.04	0.02 – 0.11	**<0.001**
Condition [Right]	1.85	0.71 – 4.86	0.210
spkr hemifield [R]	13.96	4.93 – 39.50	**<0.001**
dist to midline	7.85	3.65 – 16.86	**<0.001**
Atten	0.99	0.96 – 1.02	0.524
CenterReward	0.92	0.52 – 1.65	0.782
control pCorrect	10.61	4.67 – 24.13	**<0.001**
Condition [Right] * spkr hemifield [R]	0.17	0.05 – 0.63	**0.008**

Condition [Right] * dist to midline	0.32	0.13 – 0.81	**0.016**
spkr hemifield [R] * dist to midline	0.03	0.01 – 0.08	**<0.001**
(Condition [Right] * spkr hemifield [R]) * dist to midline	19.40	5.62 – 66.91	**<0.001**

Random Effects

σ^2	3.29
$\tau_{00\ fNum}$	0.09
ICC	0.03
$N_{\ fNum}$	2
Observations	112
Marginal R^2 / Conditional R^2	0.185 / 0.208

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.

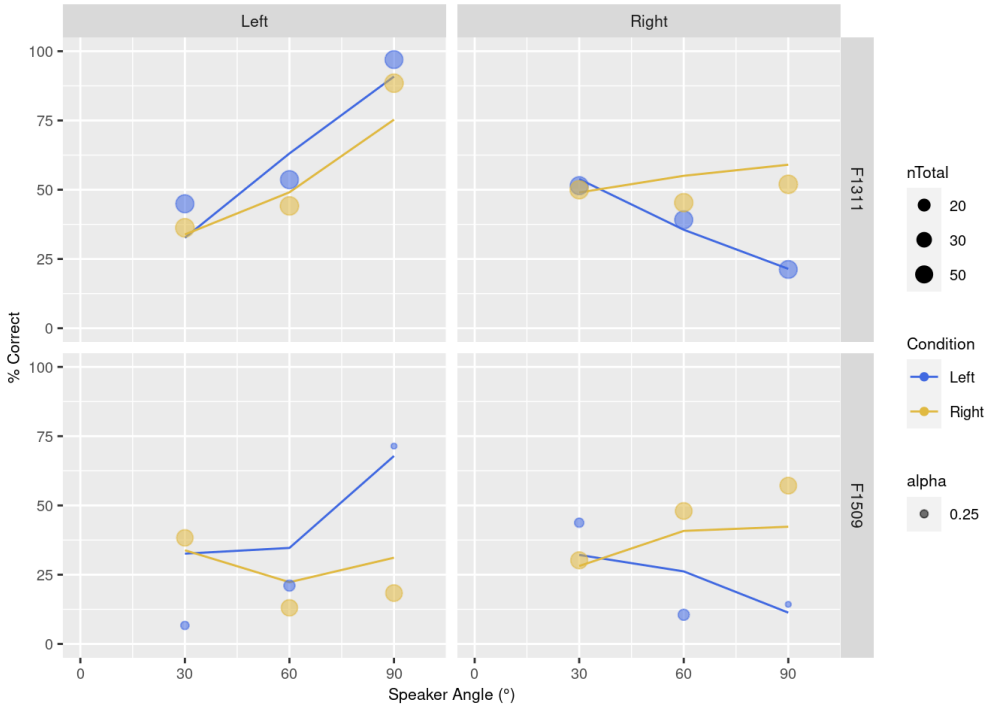
```
# Predict probability of making correct response
lr_extended$fit <- predict(LR_4, lr_extended, type="response")

# Predict the number of correct trials, using original trial counts
lr_extended <- lr_extended %>%
  mutate(
    predicted_correct = fit * nTotal,
    speaker_rad = ifelse(spkr_hemifield == 'L', -dist_to_midline, dist_to_midline)
    )

# Summarize data center reward and attenuation
lr_agg = lr_extended %>%
  group_by(fNum, Condition, dist_to_midline, spkr_hemifield) %>%
  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.

```
ggplot(                                # Plot facet grid for predicted and observed performance
  lr_agg  %>%
    mutate(dist_to_midline = dist_to_midline / pi * 180), # rad2deg
  aes(x = dist_to_midline, color=Condition)
  ) +
geom_line(
  aes(y = fit_p)
  ) +
geom_point(
  shape = 19,
  aes(
    y = pCorrect,
    size = nTotal,
    alpha = 0.25
    )
  ) +
scale_size(range = c(1, 4), trans="log10") +
scale_color_manual(values=c('#4169E1','#E0B941')) +
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(0, 100), breaks=seq(-90,90,30)) +
scale_y_continuous(name="% Correct", limits=c(0, 100)) +
facet_grid(
  fNum ~ spkr_hemifield,
  labeller = labeller(
    fNum =
      c("1311" = "F1311",
        "1509" = "F1509"),
    spkr_hemifield =
      c('L' = 'Left',
        'R' = 'Right')
  )
)
```




```
ggsave(filename = "Localization_predictions_LR2b.png", width = 10.5, height = 6.75, units = "cm", dpi=300)
```