Example of using R Markdown

An eye movement analysis of sentence reading, comparing a reader with aphasia to a neurologically healthy reader

Abstract

Mild reading difficulties are a pervasive symptom of aphasia, a language impairment common post stroke. In this study, we used eye tracking to investigate sentence reading by one person diagnosed with aphasia (PWA), compared to a neurologically healthy participant (NHI). Data were extracted from a larger project on sentence reading (published in Aphasiology online). The main aim of this study was to find out whether the eye movements of these two readers are influenced by linguistic factors of word frequency and contextual predictability. The two participants read sentences including target words that varied in word frequency and contextual predictability, and answered comprehension questions. We recorded gaze duration, total fixation duration, and first-pass regressions. Results demonstrated that the PWA had prolonged gaze and total fixation duratations and an increase of first-pass regressions compared to the NHI. Both readers were influenced by word frequency and predictability, but in different ways. Readers varied in gaze duration and first-pass regressions in particular, which may point to differences in the phase of lexical access.

Load libraries

```r
library(gdata)
library(ggplot2)
library(pastecs)
library(reshape)
library(gridExtra)
library(lme4)
library(lmerTest)
```

```r
#set working directory
setwd("~/code/r-markdown-template-example")
```

We are going to load data of the two participants from the reading study.

Open database:

```r
rawdata=read.xls("EMdataexample.xlsx",
                 na.strings = c("zero"),
                 colClasses = c(
                   'factor',  # RECORDING_SESSION_LABEL
                   'factor',  # GROUP
                   'factor',  # ID_OVERALL
                   'factor',  # ID
                   'factor',  # TRIAL_INDEX
                   'factor',  # trial_type
                   'factor',  # FREQUENCY
                   'factor',  # PREDICTABILITY
                   'factor',  # SENTENCE
                   'factor',  # ITEM
                   'factor',  # QUESTION
                   'factor',  # CRITICAL_WORD
```

```r
                   'factor',  # ACCURACY
                   'character', # SINGLE_FIXATION_DURATION
                   'character', # FIRST_FIXATION_DURATION
                   'character', # GAZE_DURATION
                   'character', # RIGHT_BOUNDED_DURATION
                   'character', # REGRESSION_PATH_DURATION
                   'character', # REREADING_DURATION
                   'character', # TOTAL_DURATION
                   'character',  # FIRST_PASS_REGRESSION
                   'character',  # FIRST_PASS_FIXATION
                   'factor',  # FIRST_PASS_MULTI_FIXATION
                   'character' #trials.fixated
                 )
)
```

```r
# rawdata
```

Create a new dataframe for analysis

```r
data <-rawdata
```

Explore the data

```r
str(data)
```

```
## 'data.frame':    112 obs. of  25 variables:
##  $ RECORDING_SESSION_LABEL  : Factor w/ 4 levels "c_11_a","c_11_b",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ GROUP                    : Factor w/ 2 levels "NHI","PWA": 1 1 1 1 1 1 1 1 1 1 ...
##  $ ID_OVERALL               : Factor w/ 2 levels "1","21": 1 1 1 1 1 1 1 1 1 1 ...
##  $ ID                       : Factor w/ 1 level "1": 1 1 1 1 1 1 1 1 1 1 ...
##  $ TRIAL_INDEX              : Factor w/ 28 levels "1","10","11",..: 1 12 22 23 24 25 26 27 28 2 ...
##  $ TRIAL_TYPE               : Factor w/ 4 levels "Experimental WP HF P",..: 4 2 3 2 4 2 3 1 2 1 ...
##  $ FREQUENCY                : Factor w/ 2 levels "high frequency",..: 2 1 2 1 2 1 2 1 1 1 ...
##  $ PREDICTABILITY           : Factor w/ 2 levels "predictable",..: 2 2 1 2 2 2 1 1 2 1 ...
##  $ SENTENCE                 : Factor w/ 56 levels "After a long day at work she forgot her keys and l
##  $ ITEM                     : Factor w/ 28 levels "bank","brewery",..: 3 22 12 1 17 18 11 25 20 27 .
##  $ QUESTION.                : Factor w/ 28 levels "Are the backpackers staying in a hostel?",..: 8 19
##  $ CRITICAL_WORD            : Factor w/ 11 levels "10","11","12",..: 10 2 8 9 11 10 3 6 3 8 ...
##  $ ACCURACY                 : Factor w/ 2 levels "correct","incorrect": 1 1 1 1 1 1 1 1 1 1 ...
##  $ SINGLE_FIXATION_DURATION : chr  "315" "153" "143" "385" ...
##  $ FIRST_FIXATION_DURATION  : chr  "315" "153" "143" "385" ...
##  $ GAZE_DURATION            : chr  "315" "153" "143" "385" ...
##  $ RIGHT_BOUNDED_DURATION   : chr  "315" "153" "143" "385" ...
##  $ REGRESSION_PATH_DURATION : chr  "315" "636" "143" "385" ...
##  $ REREADING_DURATION       : chr  "0" "483" "0" "0" ...
##  $ TOTAL_DURATION           : chr  "546" "153" "316" "385" ...
##  $ FIRST_PASS_REGRESSION    : chr  "0" "1" "0" "0" ...
##  $ FIRST_PAST_FIXATION      : chr  "1" "1" "1" "1" ...
##  $ FIRST_PASS_MULTI_FIXATION: Factor w/ 3 levels ".","0","1": 2 2 2 2 3 3 2 2 1 2 ...
##  $ trials.fixated           : chr  "1" "1" "1" "1" ...
##  $ BLINK                    : Factor w/ 1 level "no": 1 1 1 1 1 1 1 1 1 1 ...
```

```r
#str(data)
#summary(data)
#head(data[, 1:10])
#tail(data[, 1:10])
#dim(data)
```

Preparing variables we are interested in:

Create variables as numeric

```r
data$TOTAL_DURATION  <-as.numeric(data$TOTAL_DURATION)
data$GAZE_DURATION  <-as.numeric(data$GAZE_DURATION)
data$FIRST_PASS_REGRESSION <-as.numeric(data$FIRST_PASS_REGRESSION)
```

```
## Warning: NAs introduced by coercion
```

Check whether the data frame inlcudes NAs

```r
which(is.na(data$GAZE_DURATION))
```

```
##  [1]   9  12  13  16  25  26  28  29  46  47  48  52  53  59  65  67  69
## [18]  74  78  79  81  82  84  91  92  97 103 104 106 110 111
```

```r
which(is.na(data$TOTAL_DURATION))
```

```
##  [1]   9  13  16  26  28  29  46  48  52  53  69 111
```

```r
which(is.na(data$FIRST_PASS_REGRESSION))
```

```
##  [1]   9  12  13  16  25  26  28  29  46  47  48  52  53  59  65  67  69
## [18]  74  78  79  81  82  84  91  92  97 103 104 106 110 111
```

```r
which(is.na(data$FIRST_PAST_FIXATION))
```

```
## integer(0)
```

Exclude the NAs

```r
data <- data[(!is.na(data$GAZE_DURATION)),]
data <- data[(!is.na(data$TOTAL_DURATION)),]
data <- data[(!is.na(data$FIRST_PASS_REGRESSION)),]
data <- data[(!is.na(data$FIRST_PAST_FIXATION)),]
```

Check whether it worked ok

```r
which(is.na(data$GAZE_DURATION))
```

```
## integer(0)
```

```r
which(is.na(data$TOTAL_DURATION))
```

```
## integer(0)
```

```r
which(is.na(data$FIRST_PASS_REGRESSION))
```

```
## integer(0)
```

```r
which(is.na(data$FIRST_PAST_FIXATION))
```

```
## integer(0)
```

Rename GROUP as CASE - because this example dataset is restricted to the

comparison of two cases

```r
data <- rename(data, c(GROUP="CASE"))
```

Data analysis

There are four conditions (=TRIAL TYPES) in this dataset. Sentences with: * High frequency predictable words * High frequency unpredictable words * Low frequency predictable words * Low frequency unpredictable words

Independent variables are: word frequency, contextual predictability and case

Dependent variables are: gaze duration, total fixation duration, first-pass regression

We start by getting some descriptive stats, comparing the four trial types:

Gaze duration as a measure of TRIAL TYPE and CASE

```r
#by(data$GAZE_DURATION, list(data$TRIAL_TYPE, data$CASE), stat.desc, basic= FALSE)
```

Total fixation duration as a measure of TRIAL TYPE and CASE

```r
# by(data$TOTAL_DURATION, list(data$TRIAL_TYPE, data$CASE), stat.desc, basic= FALSE)
```

First-pass regression as a measure of TRIAL TYPE and CASE

```r
# by(data$FIRST_PASS_REGRESSION, list(data$TRIAL_TYPE, data$CASE), stat.desc, basic= FALSE)
```

Plotting - Eye movements independent of trial types

Gaze duration

```r
plot_gaze <- ggplot(data,aes(x=CASE, y=GAZE_DURATION, fill=CASE)) +
  stat_summary(fun.data=mean_cl_normal,position=position_dodge(0.95),geom="errorbar") +
  stat_summary(fun.y=mean,position=position_dodge(width=0.95),geom="bar") +
    ylab("Gaze duration in ms") +
    xlab("Case") +
    theme(axis.text=element_text(size=11)) +
```

```
    theme(axis.title.x=element_text(size=11)) +
    theme(axis.title.y=element_text(size=11)) +
    theme(legend.position="none") +
    ggtitle("Gaze duration NHI vs PWA") +
    theme(plot.title = element_text(size = 12, face = "bold"))
# plot_gaze
```

Total fixation duration

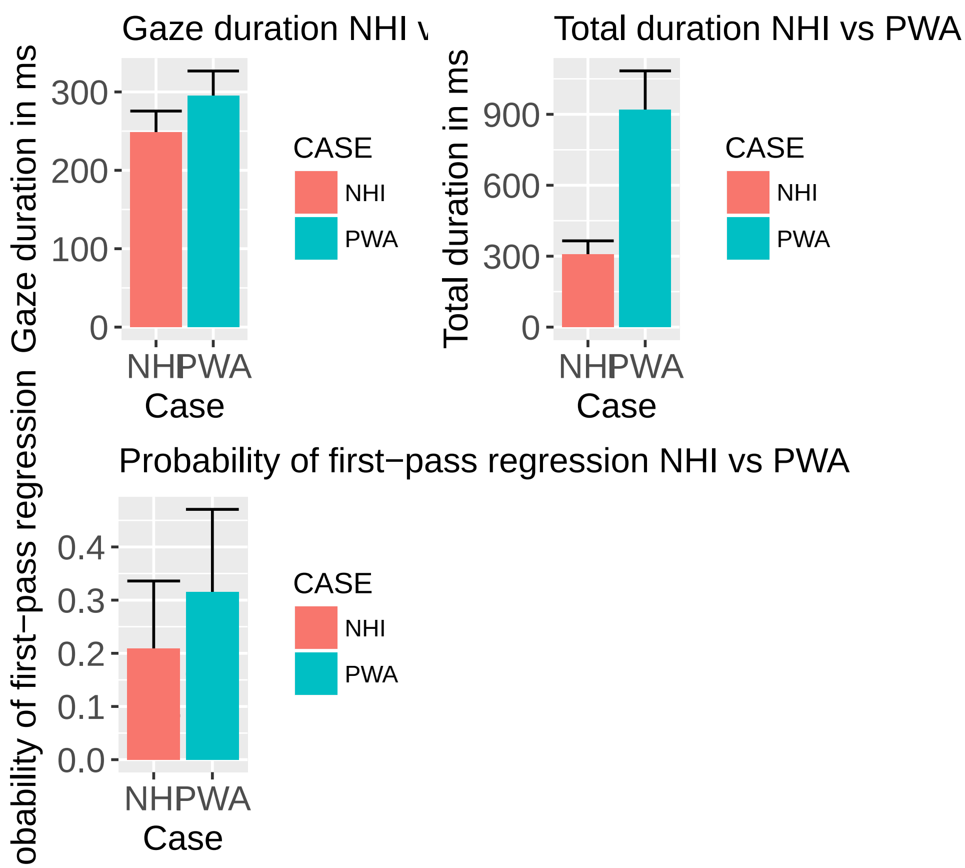
```
plot_total <- ggplot(data,aes(x=CASE, y=TOTAL_DURATION, fill=CASE)) +
  stat_summary(fun.data=mean_cl_normal,position=position_dodge(0.95),geom="errorbar") +
  stat_summary(fun.y=mean,position=position_dodge(width=0.95),geom="bar") +
    ylab("Total duration in ms") +
    xlab("Case") +
    theme(axis.text=element_text(size=11)) +
    theme(axis.title.x=element_text(size=11)) +
    theme(axis.title.y=element_text(size=11)) +
    theme(legend.position="none") +
    ggtitle("Total duration NHI vs PWA") +
    theme(plot.title = element_text(size = 12, face = "bold"))
#plot_total
```

First-pass regression

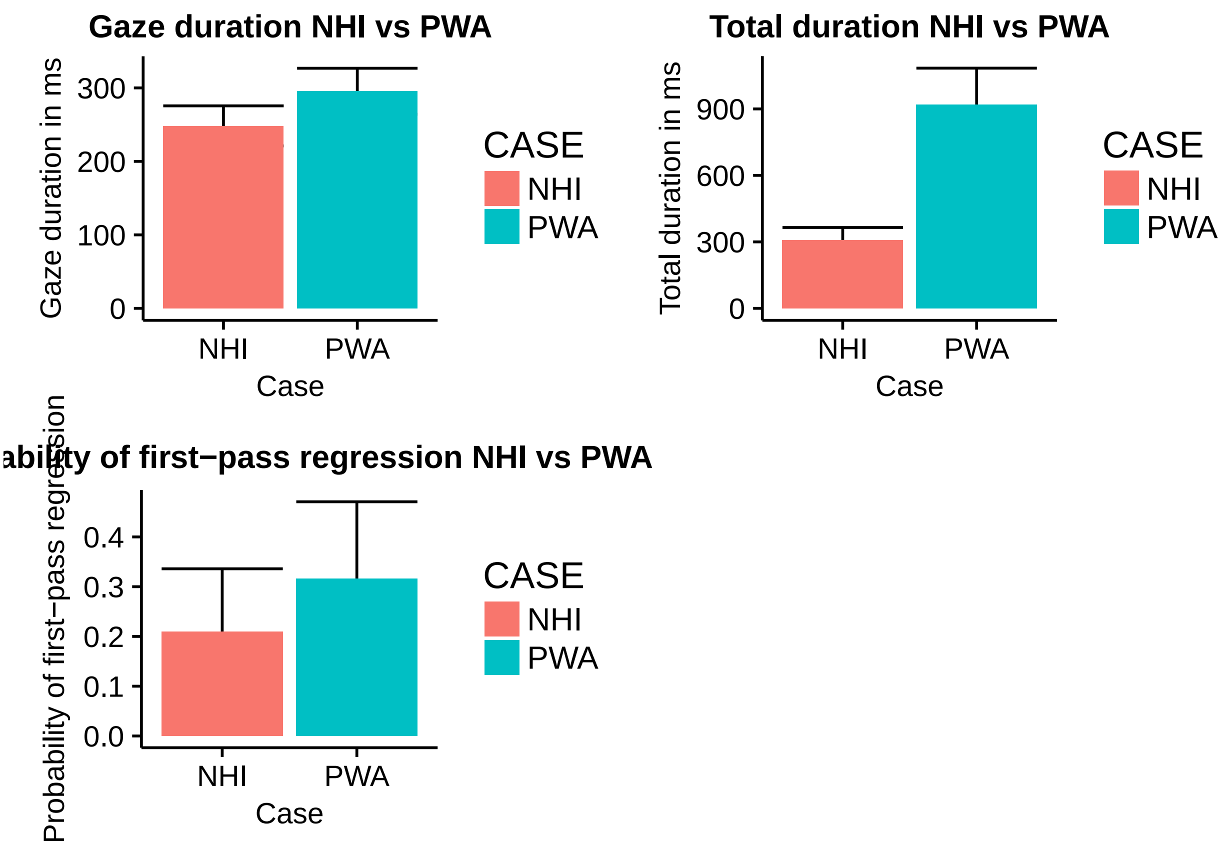
```
plot_regress_prob <- ggplot(data,aes(x=CASE, y=FIRST_PASS_REGRESSION, fill=CASE)) +
  stat_summary(fun.data=mean_cl_normal,position=position_dodge(0.95),geom="errorbar") +
  stat_summary(fun.y=mean,position=position_dodge(width=0.95),geom="bar") +
    ylab("Probability of first-pass regression") +
    xlab("Case") +
    theme(axis.text=element_text(size=11)) +
    theme(axis.title.x=element_text(size=11)) +
    theme(axis.title.y=element_text(size=11)) +
    theme(legend.position="none") +
    ggtitle("Probability of first-pass regression NHI vs PWA") +
    theme(plot.title = element_text(size = 12, face = "bold"))

#plot_regress_prob
```

```
grid.arrange(plot_gaze, plot_total, plot_regress_prob, ncol=2, respect=TRUE)
```




```
library(cowplot)
plot_grid(plot_gaze, plot_total, plot_regress_prob)
```



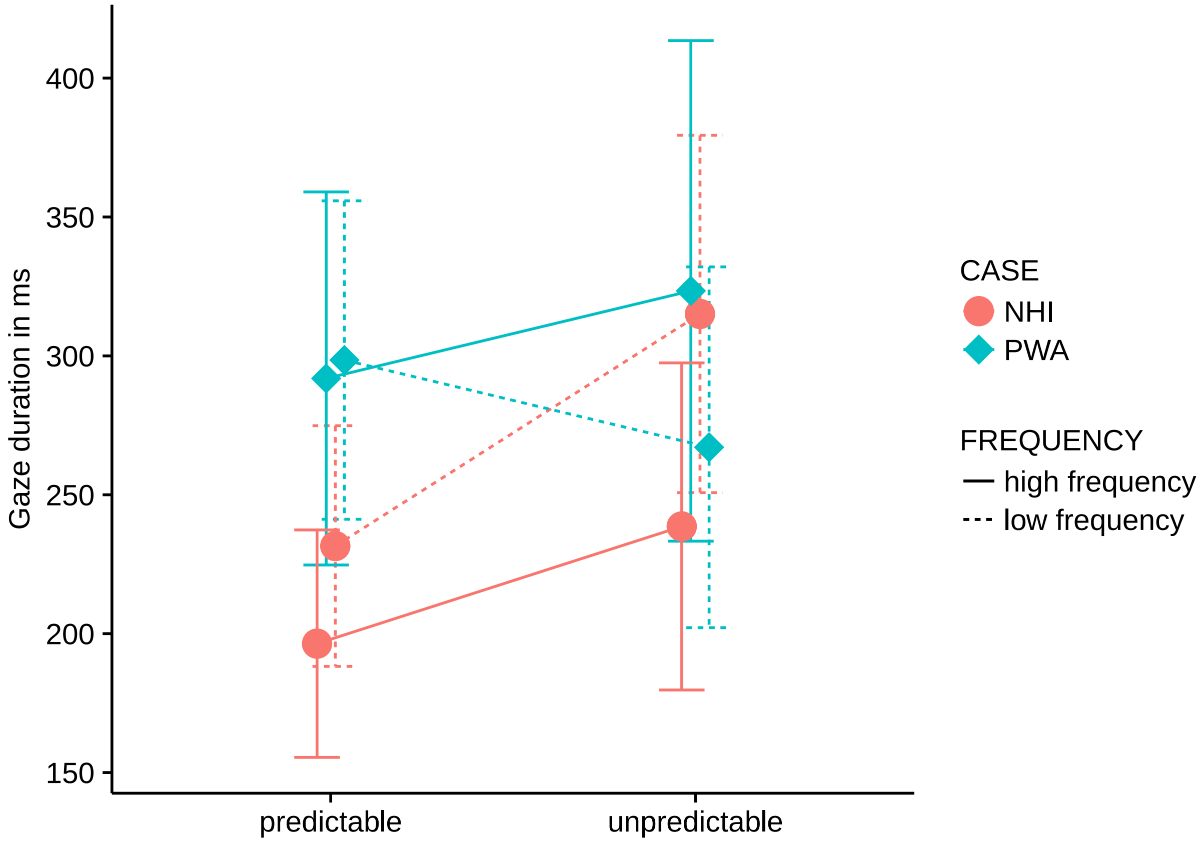
Summary:

The participant with aphasia shows an increase in reading times and in first- pass regressions.

Plotting - Eye movements as a function of trial type

Gaze duration

```
line_gaze <-
ggplot(data, aes(x=PREDICTABILITY, y=GAZE_DURATION, group=interaction(CASE, FREQUENCY), colour=CASE, li
  stat_summary(fun.data=mean_cl_normal, geom="errorbar", position=position_dodge(width=0.10), width=0.5
  stat_summary(fun.y=mean, geom="line", position=position_dodge(width=0.10)) +
  stat_summary(fun.y=mean, geom="point",position=position_dodge(width=0.10), aes(shape=CASE), size=5) +
  scale_shape_manual(values = c(16, 18)) +
  scale_x_discrete(limits=c("predictable", "unpredictable")) +
  theme (axis.text.x=element_text(colour="#000000", size=11)) +
  theme (axis.text.y=element_text(colour="#000000", size=11)) +
  theme(axis.title.y=element_text(colour="#000000", size=11)) +
  theme (axis.title.x = element_blank()) +
  scale_y_continuous(name="Gaze duration in ms") +
  theme(legend.title = element_text(size=11)) +
  theme(legend.text = element_text(size = 11)) +
  theme(legend.position="right")
 line_gaze
```



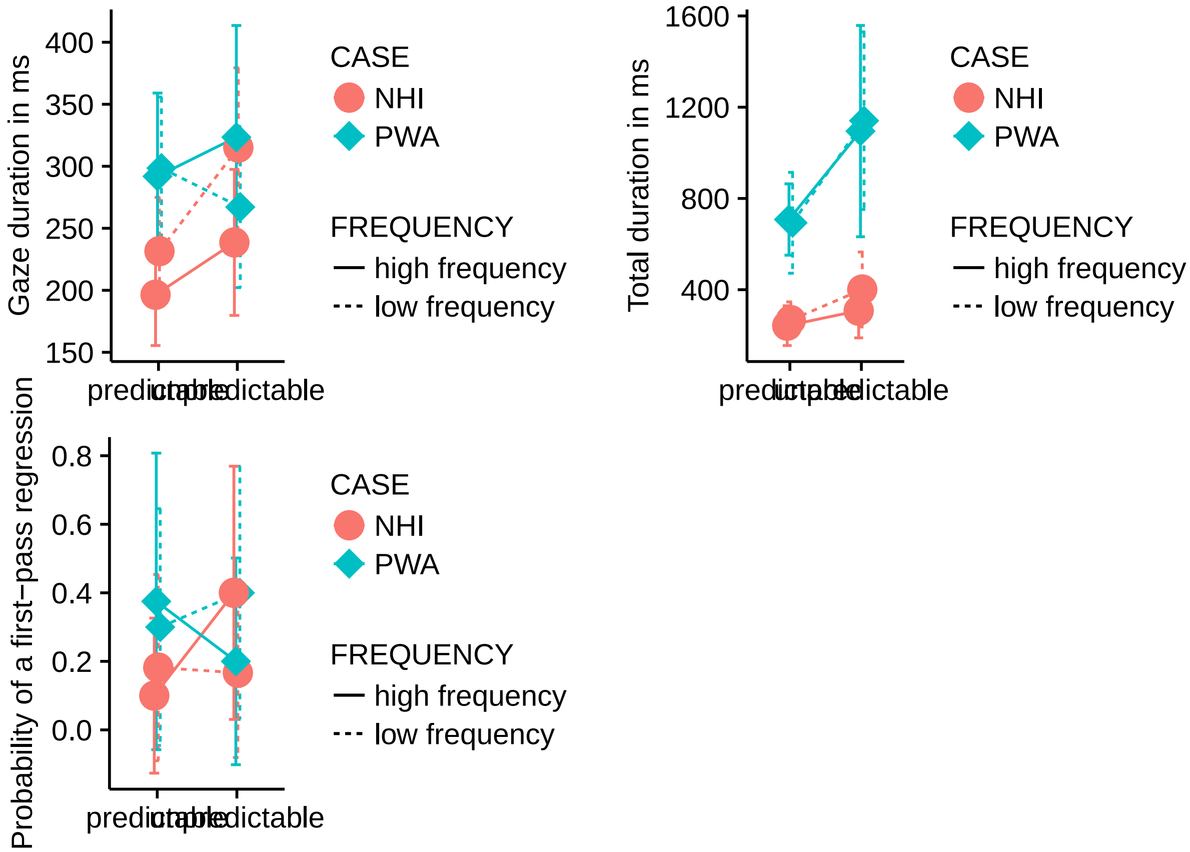
Total fixation duration

```
line_total <-
ggplot(data, aes(x=PREDICTABILITY, y=TOTAL_DURATION, group=interaction(CASE, FREQUENCY), colour=CASE, l
  stat_summary(fun.data=mean_cl_normal, geom="errorbar", position=position_dodge(width=0.10), width=0.5
  stat_summary(fun.y=mean, geom="line", position=position_dodge(width=0.10)) +
  stat_summary(fun.y=mean, geom="point",position=position_dodge(width=0.10), aes(shape=CASE), size=5) +
  scale_shape_manual(values = c(16, 18)) +
  scale_x_discrete(limits=c("predictable", "unpredictable")) +
  theme (axis.text.x=element_text(colour="#000000", size=11)) +
  theme (axis.text.y=element_text(colour="#000000", size=11)) +
  theme(axis.title.y=element_text(colour="#000000", size=11)) +
  theme (axis.title.x = element_blank()) +
  scale_y_continuous(name="Total duration in ms") +
  theme(legend.title = element_text(size=11)) +
  theme(legend.text = element_text(size = 11)) +
  theme(legend.position="right")
#line_total
```

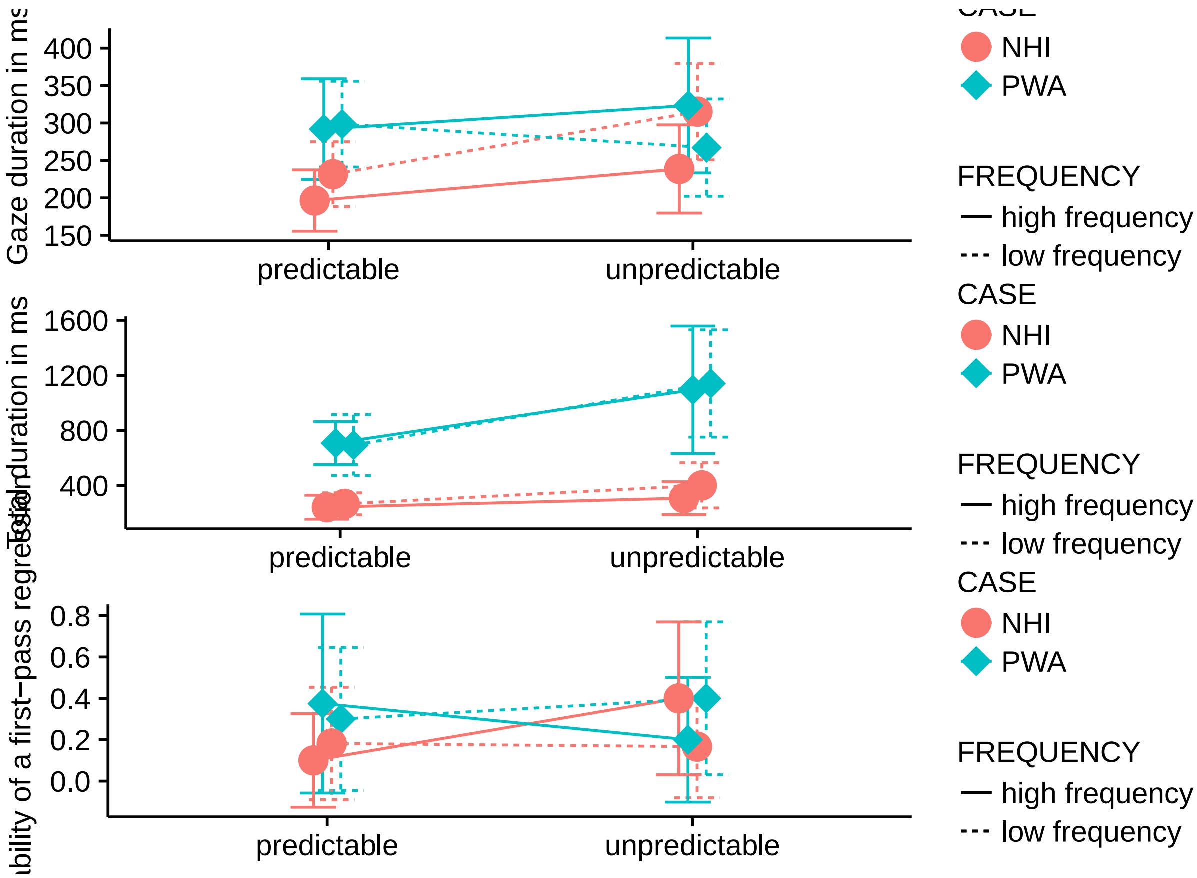
First-pass regression

```
line_regression <-
ggplot(data, aes(x=PREDICTABILITY, y=FIRST_PASS_REGRESSION, group=interaction(CASE, FREQUENCY), colour=
  stat_summary(fun.data=mean_cl_normal, geom="errorbar", position=position_dodge(width=0.10), width=0.5
  stat_summary(fun.y=mean, geom="line", position=position_dodge(width=0.10)) +
  stat_summary(fun.y=mean, geom="point",position=position_dodge(width=0.10), aes(shape=CASE), size=5) +
  scale_shape_manual(values = c(16, 18)) +
  scale_x_discrete(limits=c("predictable", "unpredictable")) +
  theme (axis.text.x=element_text(colour="#000000", size=11)) +
  theme (axis.text.y=element_text(colour="#000000", size=11)) +
  theme(axis.title.y=element_text(colour="#000000", size=11)) +
  theme (axis.title.x = element_blank()) +
  scale_y_continuous(name="Probability of a first-pass regression") +
  theme(legend.title = element_text(size=11)) +
  theme(legend.text = element_text(size = 11)) +
  theme(legend.position="right")
#line_regression
```

```
plot_grid(line_gaze, line_total, line_regression)
```




```
grid.arrange(line_gaze, line_total, line_regression, nrow=3)
```



Linear mixed model analysis of effects of word frequency and predictability

Gaze duration

```
model_simple = lmer (GAZE_DURATION ~ CASE + (1 | ITEM),
                     data=data, REML=FALSE)
##summary(model_simple)
```

```
model_a = lmer (GAZE_DURATION ~CASE+FREQUENCY + (1 | ITEM),
                     data=data, REML=FALSE)
```

```
anova(model_simple, model_a)
```

```
## Data: data
## Models:
## object: GAZE_DURATION ~ CASE + (1 | ITEM)
## ..1: GAZE_DURATION ~ CASE + FREQUENCY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## object  4 967.92 977.50 -479.96   959.92
## ..1     5 969.14 981.12 -479.57   959.14  0.78      1     0.3772
```

```
# not significant so FREQUENCY does not improve model fit
```

```
model_b = lmer (GAZE_DURATION ~CASE+PREDICTABILITY + (1 | ITEM),
                     data=data, REML=FALSE)
```

```r
anova(model_simple, model_b)
```

```
## Data: data
## Models:
## object: GAZE_DURATION ~ CASE + (1 | ITEM)
## ..1: GAZE_DURATION ~ CASE + PREDICTABILITY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  4 967.92 977.50 -479.96   959.92
## ..1     5 966.74 978.71 -478.37   956.74 3.1865      1    0.07425 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# not significant so PREDICTABILITY does not improve model fit
```

```r
# Checking for interaction between CASE and PREDICTABILITY
model_c = lmer (GAZE_DURATION~CASE + PREDICTABILITY + (1 | ITEM),
                     data=data, REML=FALSE)
model_d = lmer (GAZE_DURATION~CASE * PREDICTABILITY + (1 | ITEM),
                     data=data, REML=FALSE)
anova(model_c, model_d) # not significant so no interaction between CASE and PREDICTABIITY
```

```
## Data: data
## Models:
## object: GAZE_DURATION ~ CASE + PREDICTABILITY + (1 | ITEM)
## ..1: GAZE_DURATION ~ CASE * PREDICTABILITY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  5 966.74 978.71 -478.37   956.74
## ..1     6 965.84 980.21 -476.92   953.84 2.8975      1    0.08872 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Checking for interaction between CASE and FREQUENCY
model_e = lmer (GAZE_DURATION~CASE + FREQUENCY + (1 | ITEM),
                     data=data, REML=FALSE)
model_f = lmer (GAZE_DURATION~CASE * FREQUENCY + (1 | ITEM),
                     data=data, REML=FALSE)
anova(model_e, model_f) # there is a significant interaction between CASE and FREQUENCY
```

```
## Data: data
## Models:
## object: GAZE_DURATION ~ CASE + FREQUENCY + (1 | ITEM)
## ..1: GAZE_DURATION ~ CASE * FREQUENCY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  5 969.14 981.12 -479.57   959.14
## ..1     6 966.64 981.01 -477.32   954.64 4.5027      1    0.03384 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Total duration

```r
model_TD_simple = lmer (TOTAL_DURATION ~CASE + (1 | ITEM),
                       data=data, REML=FALSE)
# summary(model_TD_simple)
```

```r
model_TD_a = lmer (TOTAL_DURATION ~CASE+FREQUENCY + (1 | ITEM),
                       data=data, REML=FALSE)
```

```r
anova(model_TD_simple, model_TD_a)
```

```
## Data: data
## Models:
## object: TOTAL_DURATION ~ CASE + (1 | ITEM)
## ..1: TOTAL_DURATION ~ CASE + FREQUENCY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  4 1192.5 1202.1 -592.24   1184.5
## ..1     5 1194.3 1206.3 -592.18   1184.3 0.1373      1      0.711
```

```r
# not significant so FREQUENCY does not improve model fit
```

```r
model_TD_b = lmer (TOTAL_DURATION ~CASE+PREDICTABILITY + (1 | ITEM),
                       data=data, REML=FALSE)
```

```r
anova(model_TD_simple, model_TD_b)
```

```
## Data: data
## Models:
## object: TOTAL_DURATION ~ CASE + (1 | ITEM)
## ..1: TOTAL_DURATION ~ CASE + PREDICTABILITY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  4 1192.5 1202.1 -592.24   1184.5
## ..1     5 1184.1 1196.0 -587.03   1174.1 10.433      1   0.001238 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# significant so PREDICTABILITY does improve model fit
```

```r
# Checking for interaction between CASE and PREDICTABILITY
model_TD_c = lmer (TOTAL_DURATION~CASE + PREDICTABILITY + (1 | ITEM),
                     data=data, REML=FALSE)
model_TD_d = lmer (TOTAL_DURATION~CASE * PREDICTABILITY + (1 | ITEM),
                     data=data, REML=FALSE)
anova(model_TD_c, model_TD_d) # not significant so no interaction between CASE and PREDICTABIITY
```

```
## Data: data
## Models:
## object: TOTAL_DURATION ~ CASE + PREDICTABILITY + (1 | ITEM)
## ..1: TOTAL_DURATION ~ CASE * PREDICTABILITY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  5 1184.1 1196.0 -587.03   1174.1
## ..1     6 1181.3 1195.6 -584.63   1169.3 4.7953      1    0.02854 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Checking for interaction between CASE and FREQUENCY
model_TD_e = lmer (TOTAL_DURATION~CASE + FREQUENCY + (1 | ITEM),
                     data=data, REML=FALSE)
model_TD_f = lmer (TOTAL_DURATION~CASE * FREQUENCY + (1 | ITEM),
                     data=data, REML=FALSE)
anova(model_TD_e, model_TD_f) # no significant interaction between CASE and FREQUENCY
```

```
## Data: data
```

```
## Models:
## object: TOTAL_DURATION ~ CASE + FREQUENCY + (1 | ITEM)
## ..1: TOTAL_DURATION ~ CASE * FREQUENCY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  5 1194.3 1206.3 -592.18   1184.3
## ..1     6 1196.2 1210.5 -592.09   1184.2 0.1742      1     0.6764
```

First-pass regression

```
model_R_simple = lmer (FIRST_PASS_REGRESSION ~CASE + (1 | ITEM),
                       data=data, REML=FALSE)
# summary(model_R_simple)
```

```
model_R_a = lmer (FIRST_PASS_REGRESSION ~CASE+FREQUENCY + (1 | ITEM),
                       data=data, REML=FALSE)
```

```
anova(model_R_simple, model_R_a)
```

```
## Data: data
## Models:
## object: FIRST_PASS_REGRESSION ~ CASE + (1 | ITEM)
## ..1: FIRST_PASS_REGRESSION ~ CASE + FREQUENCY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  4 102.98 112.56 -47.492   94.984
## ..1     5 104.98 116.95 -47.491   94.981 0.0029      1     0.9574
```

```
# not significant so FREQUENCY does not improve model fit
```

```
model_R_b = lmer (FIRST_PASS_REGRESSION ~CASE+PREDICTABILITY + (1 | ITEM),
                       data=data, REML=FALSE)
```

```
anova(model_R_simple, model_R_b)
```

```
## Data: data
## Models:
## object: FIRST_PASS_REGRESSION ~ CASE + (1 | ITEM)
## ..1: FIRST_PASS_REGRESSION ~ CASE + PREDICTABILITY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  4 102.98 112.56 -47.492   94.984
## ..1     5 104.70 116.67 -47.349   94.699 0.2853      1     0.5932
```

```
# not significant so PREDICTABILITY does not improve model fit
```

```
# Checking for interaction between CASE and PREDICTABILITY
model_R_c = lmer (FIRST_PASS_REGRESSION~CASE + PREDICTABILITY + (1 | ITEM),
                     data=data, REML=FALSE)
model_R_d = lmer (FIRST_PASS_REGRESSION~CASE * PREDICTABILITY + (1 | ITEM),
                     data=data, REML=FALSE)
anova(model_R_c, model_R_d) #  not significant so there is no interaction between CASE and PREDICTABIIT
```

```
## Data: data
## Models:
## object: FIRST_PASS_REGRESSION ~ CASE + PREDICTABILITY + (1 | ITEM)
## ..1: FIRST_PASS_REGRESSION ~ CASE * PREDICTABILITY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  5 104.70 116.67 -47.349   94.699
```

```
## ..1     6 105.97 120.33 -46.982   93.965 0.7338      1     0.3917
```

```r
# Checking for interaction between CASE and FREQUENCY
model_R_e = lmer (FIRST_PASS_REGRESSION~CASE + FREQUENCY + (1 | ITEM),
                     data=data, REML=FALSE)
model_R_f = lmer (FIRST_PASS_REGRESSION~CASE * FREQUENCY + (1 | ITEM),
                     data=data, REML=FALSE)
anova(model_R_e, model_R_f) # no significant interaction between CASE and FREQUENCY
```

```
## Data: data
## Models:
## object: FIRST_PASS_REGRESSION ~ CASE + FREQUENCY + (1 | ITEM)
## ..1: FIRST_PASS_REGRESSION ~ CASE * FREQUENCY + (1 | ITEM)
##        Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## object  5 104.98 116.95 -47.491   94.981
## ..1     6 106.41 120.78 -47.207   94.413 0.5678      1     0.4512
```

Summary:

Eye movements by both participants are influenced by word frequency and contextual predictability, but in inconsistent ways. The neurologically healthy participant demonstrates a word frequency effect in the predicted direction for gaze duration (increase in gaze duration for low frequency words), and a predictability effect in the expected direction for total duration (prolonged total fixation durations on unpredictable words). The participant with aphasia showed a word frequency effect for gaze duration that was in the non-predicted direction (longer gaze duration for high frequency words), but a predictability effect for total fixation duration in the expected direction and in parallel to the neurologically healthy participant. Both participants seemed to be differently affected by word frequency and predictability with respect to first-pass regressions. The neurologically healthy particiapnts was more likely to regress out of high frequency words if they were unpredictable than low frequency words. The participant with aphasia, however, regressed more out of unpredictable low frequency words than unpredictable high frequency words. However, the models did not find that this difference between participants was significant.