sentence\_level\_analysis

2024-06-03

Version: Rtudio: 2023.12.0 Build 369 R:4.4.0 lme4:1.1-35.1 lattice:0.22-6 MASS:7.3-60.2 plyr:1.8.9 plotrix:3.8-4

Load the packages that we used

rm(list=ls()) # clear objects from the workspace  
datafile <- read.csv("datafile.csv", sep = ",", dec = ".") # open the datafile  
colnames(datafile) #retrieve the column names of the datafile

## [1] "Participant" "Item"   
## [3] "Age" "Predictability"   
## [5] "Preview" "Age\_Predictability"   
## [7] "Age\_Preview" "Predictability\_Preview"   
## [9] "Condition" "TRT"   
## [11] "IA\_SKIP" "FFD"   
## [13] "SFD" "GD"   
## [15] "RO" "RI"   
## [17] "sentence.reading.time" "average.fixation.duration"   
## [19] "number.of.fixations" "number.of.regressions"   
## [21] "average.forwary.saccade.length"

nrow(datafile) #the number of observations of the datafile

## [1] 6811

datafile$pp = datafile[,1]  
datafile$Item = datafile[,2]  
datafile$Age = datafile[,3]  
# Let's make sure all the variables are from the correct class  
datafile$Item = as.factor(datafile$Item)  
datafile$Age = factor(datafile$Age, levels = c("Y","O"))

# MEAN AND STANDARD ERROR

mean.matrix = tapply(datafile$sentence.reading.time, list(datafile$pp, datafile$Age), mean, na.rm = T)  
sd.matrix = tapply(datafile$sentence.reading.time, list(datafile$pp, datafile$Age), sd, na.rm = T)  
se.matrix = tapply(datafile$sentence.reading.time, list(datafile$pp, datafile$Age), std.error, na.rm = T)  
grand.mean=apply(mean.matrix, 2, mean, na.rm = T)  
grand.sd=apply(sd.matrix,2,mean, na.rm = T)  
grand.se=apply(se.matrix,2, mean, na.rm = T)  
sentence.reading.time <- data.frame(grand.mean,grand.sd,grand.se)  
sentence.reading.time

## grand.mean grand.sd grand.se  
## Y 3541.001 1241.463 150.7081  
## O 7432.477 3036.095 352.4032

mean.matrix = tapply(datafile$average.fixation.duration, list(datafile$pp, datafile$Age), mean, na.rm = T)  
sd.matrix = tapply(datafile$average.fixation.duration, list(datafile$pp, datafile$Age), sd, na.rm = T)  
se.matrix = tapply(datafile$average.fixation.duration, list(datafile$pp, datafile$Age), std.error, na.rm = T)  
grand.mean=apply(mean.matrix, 2, mean, na.rm = T)  
grand.sd=apply(sd.matrix,2,mean, na.rm = T)  
grand.se=apply(se.matrix,2, mean, na.rm = T)  
average.fixation.duration <- data.frame(grand.mean,grand.sd,grand.se)  
average.fixation.duration

## grand.mean grand.sd grand.se  
## Y 216.3927 26.31561 3.202602  
## O 247.3246 22.16341 2.593033

mean.matrix = tapply(datafile$number.of.fixations, list(datafile$pp, datafile$Age), mean, na.rm = T)  
sd.matrix = tapply(datafile$number.of.fixations, list(datafile$pp, datafile$Age), sd, na.rm = T)  
se.matrix = tapply(datafile$number.of.fixations, list(datafile$pp, datafile$Age), std.error, na.rm = T)  
grand.mean=apply(mean.matrix, 2, mean, na.rm = T)  
grand.sd=apply(sd.matrix,2,mean, na.rm = T)  
grand.se=apply(se.matrix,2, mean, na.rm = T)  
number.of.fixations <- data.frame(grand.mean,grand.sd,grand.se)  
number.of.fixations

## grand.mean grand.sd grand.se  
## Y 9.856161 2.894969 0.3500243  
## O 18.951029 4.967301 0.5788250

mean.matrix = tapply(datafile$number.of.regressions, list(datafile$pp, datafile$Age), mean, na.rm = T)  
sd.matrix = tapply(datafile$number.of.regressions, list(datafile$pp, datafile$Age), sd, na.rm = T)  
se.matrix = tapply(datafile$number.of.regressions, list(datafile$pp, datafile$Age), std.error, na.rm = T)  
grand.mean=apply(mean.matrix, 2, mean, na.rm = T)  
grand.sd=apply(sd.matrix,2,mean, na.rm = T)  
grand.se=apply(se.matrix,2, mean, na.rm = T)  
number.of.regressions<- data.frame(grand.mean,grand.sd,grand.se)  
number.of.regressions

## grand.mean grand.sd grand.se  
## Y 2.360253 1.285127 0.1708801  
## O 4.609321 2.197991 0.2591854

mean.matrix = tapply(datafile$average.forwary.saccade.length, list(datafile$pp, datafile$Age), mean, na.rm = T)  
sd.matrix = tapply(datafile$average.forwary.saccade.length, list(datafile$pp, datafile$Age), sd, na.rm = T)  
se.matrix = tapply(datafile$average.forwary.saccade.length, list(datafile$pp, datafile$Age), std.error, na.rm = T)  
grand.mean=apply(mean.matrix, 2, mean, na.rm = T)  
grand.sd=apply(sd.matrix,2,mean, na.rm = T)  
grand.se=apply(se.matrix,2, mean, na.rm = T)  
average.forwary.saccade.length<- data.frame(grand.mean,grand.sd,grand.se)  
average.forwary.saccade.length

## grand.mean grand.sd grand.se  
## Y 2.631373 0.6041630 0.07414441  
## O 2.017899 0.3336016 0.03910925

Sliding contrasts

contrasts(datafile$Age) <- contr.sdif(2)

# sentence.reading.time

measure = "sentence.reading.time"  
datafile$depvar = log(datafile[,measure])  
datafile$depvar = as.numeric(datafile$depvar)  
model = lmer(depvar~Age+(1|pp)+(1|Item),datafile)  
summary(model, corr = FALSE)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: depvar ~ Age + (1 | pp) + (1 | Item)  
## Data: datafile  
##   
## REML criterion at convergence: 452.3  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.2312 -0.6605 -0.0727 0.5667 6.8653   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## pp (Intercept) 0.05991 0.2448   
## Item (Intercept) 0.06058 0.2461   
## Residual 0.05549 0.2356   
## Number of obs: 6811, groups: pp, 96; Item, 90  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 8.44557 0.03671 230.1  
## Age2-1 0.73215 0.05194 14.1

confint(model, method="Wald")

## 2.5 % 97.5 %  
## .sig01 NA NA  
## .sig02 NA NA  
## .sigma NA NA  
## (Intercept) 8.373619 8.5175180  
## Age2-1 0.630346 0.8339478

# average.fixation.duration

measure = "average.fixation.duration"  
datafile$depvar = log(datafile[,measure])  
datafile$depvar = as.numeric(datafile$depvar)  
model = lmer(depvar~Age+(1|pp)+(1|Item),datafile)  
summary(model, corr = FALSE)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: depvar ~ Age + (1 | pp) + (1 | Item)  
## Data: datafile  
##   
## REML criterion at convergence: -10578.7  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -4.0014 -0.6623 -0.0396 0.5939 4.9945   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## pp (Intercept) 0.0115594 0.10751   
## Item (Intercept) 0.0003821 0.01955   
## Residual 0.0114554 0.10703   
## Number of obs: 6811, groups: pp, 96; Item, 90  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 5.43215 0.01160 468.399  
## Age2-1 0.13556 0.02283 5.939

confint(model, method="Wald")

## 2.5 % 97.5 %  
## .sig01 NA NA  
## .sig02 NA NA  
## .sigma NA NA  
## (Intercept) 5.40942045 5.4548809  
## Age2-1 0.09081841 0.1802925

# number.of.fixations

measure = "number.of.fixations"  
datafile$depvar = log(datafile[,measure])  
datafile$depvar = as.numeric(datafile$depvar)  
model = lmer(depvar~Age+(1|pp)+(1|Item),datafile)  
summary(model, corr = FALSE)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: depvar ~ Age + (1 | pp) + (1 | Item)  
## Data: datafile  
##   
## REML criterion at convergence: 611.4  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.4680 -0.6790 -0.0630 0.6033 5.5277   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## pp (Intercept) 0.06498 0.2549   
## Item (Intercept) 0.01179 0.1086   
## Residual 0.05800 0.2408   
## Number of obs: 6811, groups: pp, 96; Item, 90  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 2.54373 0.02936 86.63  
## Age2-1 0.65657 0.05408 12.14

confint(model, method="Wald")

## 2.5 % 97.5 %  
## .sig01 NA NA  
## .sig02 NA NA  
## .sigma NA NA  
## (Intercept) 2.4861755 2.6012791  
## Age2-1 0.5505792 0.7625708

# number.of.regressions

measure = "number.of.regressions"  
datafile$depvar = log(datafile[,measure])  
datafile$depvar = as.numeric(datafile$depvar)  
model = lmer(depvar~Age+(1|pp)+(1|Item),datafile)  
summary(model, corr = FALSE)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: depvar ~ Age + (1 | pp) + (1 | Item)  
## Data: datafile  
##   
## REML criterion at convergence: 9360.6  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.3951 -0.7366 0.0630 0.7186 3.9023   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## pp (Intercept) 0.13368 0.3656   
## Item (Intercept) 0.01759 0.1326   
## Residual 0.25780 0.5077   
## Number of obs: 5984, groups: pp, 96; Item, 90  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 0.98961 0.04157 23.803  
## Age2-1 0.66775 0.07831 8.527

confint(model, method="Wald")

## 2.5 % 97.5 %  
## .sig01 NA NA  
## .sig02 NA NA  
## .sigma NA NA  
## (Intercept) 0.9081223 1.0710914  
## Age2-1 0.5142733 0.8212248

# average.forwary.saccade.length

measure = "average.forwary.saccade.length"  
datafile$depvar = log(datafile[,measure])  
datafile$depvar = as.numeric(datafile$depvar)  
model = lmer(depvar~Age+(1|pp)+(1|Item),datafile)  
summary(model, corr = FALSE)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: depvar ~ Age + (1 | pp) + (1 | Item)  
## Data: datafile  
##   
## REML criterion at convergence: -2977.4  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -3.7116 -0.6239 -0.0478 0.5723 5.0903   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## pp (Intercept) 0.045463 0.21322   
## Item (Intercept) 0.005283 0.07268   
## Residual 0.034274 0.18513   
## Number of obs: 6811, groups: pp, 96; Item, 90  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) 0.79442 0.02386 33.296  
## Age2-1 -0.24325 0.04519 -5.383

confint(model, method="Wald")

## 2.5 % 97.5 %  
## .sig01 NA NA  
## .sig02 NA NA  
## .sigma NA NA  
## (Intercept) 0.7476549 0.8411816  
## Age2-1 -0.3318223 -0.1546748