

Canonicity effects in self-paced listening

Doro

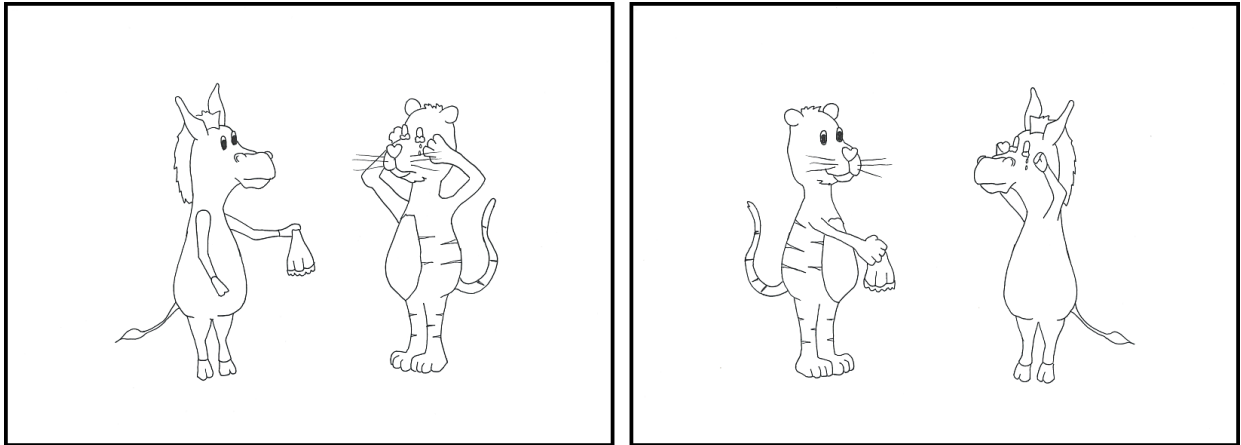
6/23/2021

```
library(tidyverse)
library(lme4)
library(car)
```

Data

This dataset includes data of 50 language unimpaired control participants and 21 individuals with aphasia, all German native speakers. Participants carried out a self-paced listening experiment. While listening to the sentence, two pictures were displayed. After sentence end, the picture that matched the sentence best had to be selected. We manipulated the word order of the sentences: Either, the subject preceded the object (canonical) or the subject followed the object (non-canonical).

For example, the pictures displayed a donkey comforting a tiger and a tiger comforting a donkey (see below). For the canonical sentence *Hier tröstet **der**_{nom} Tiger gerade **den**_{acc} Esel* “Here comforts **the**_{nom} tiger just **the**_{acc} donkey” the picture displaying the tiger comforting the donkey had to be selected (right picture). For the non-canonical sentence *Hier tröstet **den**_{acc} Tiger gerade **der**_{nom} Esel* “Here comforts **the**_{acc} tiger just **the**_{nom} donkey” the picture displaying the donkey comforting the tiger had to be selected (left picture).



Description of the columns:

- subj: subject number
- item: number of the item (number of items: n=20)
- condition: a (canonical) and b (non-canonical), we did not use a latin square design, i.e., all participants saw both conditions of all items
- region: region in the sentence (0, 1, 2, 3, question)
 - 0: sentence onset (“hier”)
 - 1: verb (“tröstet”)
 - 2: NP1 (“der Tiger”)

- 3: adverb (“gerade”)
- question: NP2 + picture selection (“den Esel” + picture selection)
- subj_status: healthy control (HC) or individual with aphasia (IWA)
- sentence: the respective auditory presented sentence
- word: the respective word of the sentence
- rt: listening time for the respective region
- acc: picture selection accuracy
- age, sex, years of education: demographic data of the participant
- aphasia_type: Aachen aphasia test syndrome, NA for control participants
- digit_symbol_substitution: WAIS subtest, number of correct items in two minutes
- digit span, block span: WMS-R subtests, forward and backward raw scores and longest span
- picture span: working memory test, forward and backward raw scores and longest span
- lexdec: lexical decision task, mean reaction times in the task

```
dat<-read.table("SPL_decl.csv", sep=",", header=TRUE)
```

Basic questions

1. Do we see faster listening times for canonical compared to non-canonical sentences in the critical region (region 2), i.e., do we see a canonicity effect?
2. Does the canonicity effect interact with participant group?

```
datcrit<-subset(dat, region==2)
```

```
# quick plot of the canonicity effect
```

```
datcrit%>%
```

```
  group_by(subj_status, condition)%>%
```

```
  dplyr::summarize(Mrt=mean(rt), SDrt=sd(rt), SErt=SDrt/sqrt(n()))%>%
```

```
  mutate(condition=ifelse(condition=="a", "canonical", "non-canonical"),
```

```
         Group=ifelse(subj_status=="HC", "Control", "Aphasia"))%>%
```

```
  ggplot(aes(x=condition, y=Mrt, color=Group, group=Group))+
```

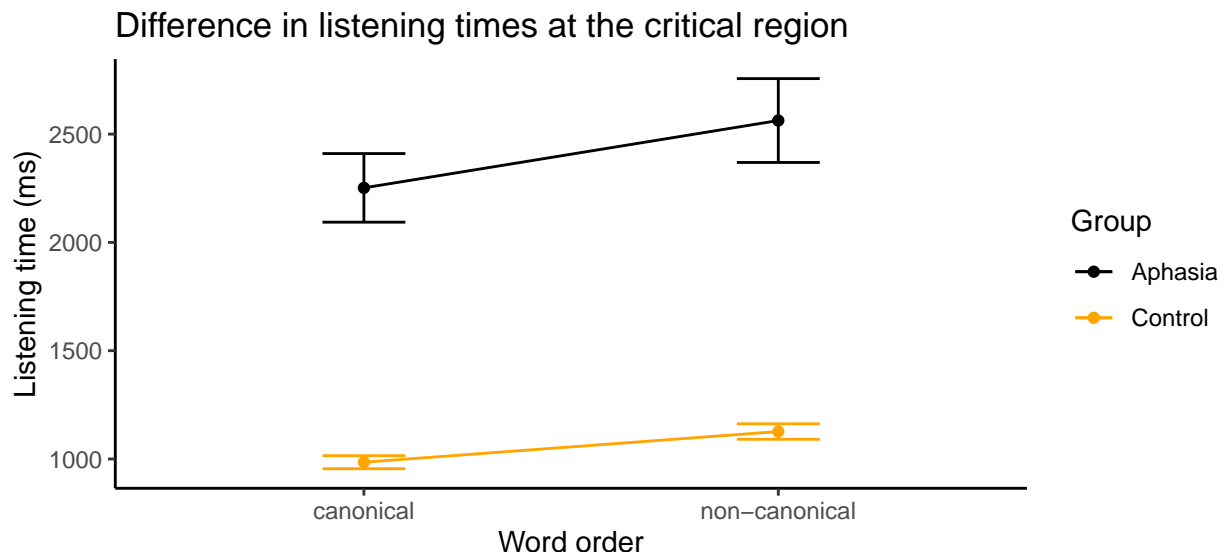
```
  geom_point()+geom_line()+
```

```
  geom_errorbar(aes(ymin=Mrt-2*SErt, ymax=Mrt+2*SErt), width=0.2)+
```

```
  theme_classic()+xlab("Word order")+ylab("Listening time (ms)")+
```

```
  scale_color_manual(values=c("black", "orange"))+
```

```
  ggtitle("Difference in listening times at the critical region")
```



```

# contrast coding
# I expect faster RTs for canonical trials
datcrit$cond<-ifelse(datcrit$condition=="a", -1, 1)
# I expect faster RTs for controls
datcrit$group<-ifelse(datcrit$subj_status=="HC", -1, 1)
datcrit$condxxgroup<-datcrit$cond*datcrit$group

# sanity checks
# there should be 50 observations for controls and 21 for individuals with aphasia
xtabs(~cond+item+subj_status,datcrit)

```

```

## , , subj_status = HC
##
##      item
## cond 10retest 10test 1retest 1test 2retest 2test 3retest 3test 4retest 4test
##  -1          50     50      50    50      50     50      49     50      50    50
##   1          50     50      50    50      50     50      50     50      49    50
##      item
## cond 5retest 5test 6retest 6test 7retest 7test 8retest 8test 9retest 9test
##  -1          50     50      50    50      50     50      50     50      50    50
##   1          49     50      50    50      50     50      50     50      50    50
##

```

```

## , , subj_status = IWA
##
##      item
## cond 10retest 10test 1retest 1test 2retest 2test 3retest 3test 4retest 4test
##  -1          21     21      21    21      21     21      21     21      21    21
##   1          21     21      21    21      21     21      21     21      21    21
##      item
## cond 5retest 5test 6retest 6test 7retest 7test 8retest 8test 9retest 9test
##  -1          21     21      21    21      21     21      21     21      21    21
##   1          21     21      21    21      21     21      21     21      21    21
##

```

```

xtabs(~group+item+condition, datcrit)

```

```

## , , condition = a
##
##      item
## group 10retest 10test 1retest 1test 2retest 2test 3retest 3test 4retest 4test
##  -1          50     50      50    50      50     50      49     50      50    50
##   1          21     21      21    21      21     21      21     21      21    21
##      item
## group 5retest 5test 6retest 6test 7retest 7test 8retest 8test 9retest 9test
##  -1          50     50      50    50      50     50      50     50      50    50
##   1          21     21      21    21      21     21      21     21      21    21
##

```

```

## , , condition = b
##
##      item
## group 10retest 10test 1retest 1test 2retest 2test 3retest 3test 4retest 4test
##  -1          50     50      50    50      50     50      50     50      49    50
##   1          21     21      21    21      21     21      21     21      21    21
##      item

```

```
## group 5retest 5test 6retest 6test 7retest 7test 8retest 8test 9retest 9test
##      -1      49      50      50      50      50      50      50      50      50
##      1      21      21      21      21      21      21      21      21      21

#the missing observations are correct (data collection error in subj 40)

m1<-lmer(log(rt)~cond+group+condxxgroup+(1+group+cond+condxxgroup||item)+
        (1+cond|subj), data = datcrit)
summary(m1)

## Linear mixed model fit by REML ['lmerMod']
## Formula: log(rt) ~ cond + group + condxxgroup + ((1 | item) + (0 + group |
##      item) + (0 + cond | item) + (0 + condxxgroup | item)) + (1 +
##      cond | subj)
##      Data: datcrit
##
## REML criterion at convergence: 1706.4
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.5154 -0.5278 -0.0686  0.3976  6.0031
##
## Random effects:
##      Groups      Name      Variance Std.Dev. Corr
##      subj      (Intercept) 0.1196003 0.34583
##              cond          0.0020381 0.04514  0.42
##      item      condxxgroup 0.0003273 0.01809
##      item.1     cond        0.0001649 0.01284
##      item.2     group        0.0002867 0.01693
##      item.3     (Intercept) 0.0012335 0.03512
##      Residual              0.0935281 0.30582
## Number of obs: 2837, groups:  subj, 71; item, 20
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  7.232496   0.046076 156.968
## cond         0.053563   0.009069   5.906
## group        0.360566   0.045560   7.914
## condxxgroup -0.006794   0.009506  -0.715
##
## Correlation of Fixed Effects:
##              (Intr) cond  group
## cond         0.267
## group        0.401  0.110
## condxxgroup  0.104  0.350  0.257

#qqPlot(residuals(m1))
```

My modeling issue

I am interested in the question whether the different cognitive factors (age, years of education, working memory scores, lexical decision times, score in the digit symbol substitution test) influence the canonicity effect in the critical region in the two participant groups. A minimal example for the factor age is given below.

```

# Centering
datcrit$c_age<-scale(datcrit$age, center = TRUE, scale = FALSE)
datcrit$c_edu<-scale(datcrit$years_of_education, center = TRUE, scale = FALSE)
datcrit$c_ds<-scale(datcrit$digit_symbol_substitution, center = TRUE, scale = FALSE)
datcrit$c_lexdec<-scale(datcrit$lexdec, center = TRUE, scale = FALSE)
# for working memory, I will use the mean of the longest spans in the
# digit, block and picture span tasks
datcrit$wm_composite_score<-rowMeans(
  datcrit[,which(grepl("longest_span", colnames(datcrit)))], na.rm = TRUE)
datcritc_wm<-scale(datcrit$wm_composite_score, center = TRUE, scale=FALSE)

# Age
m2<-lmer(log(rt)~(cond+group+condxxgroup)*c_age+(1+group+cond+condxxgroup||item)+
  (1+cond|subj), data = datcrit)
summary(m2)

```

```

## Linear mixed model fit by REML ['lmerMod']
## Formula: log(rt) ~ (cond + group + condxxgroup) * c_age + ((1 | item) +
##      (0 + group | item) + (0 + cond | item) + (0 + condxxgroup |
##      item)) + (1 + cond | subj)
##      Data: datcrit
##
## REML criterion at convergence: 1720.4
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.5106 -0.5267 -0.0650  0.3986  5.9394
##
## Random effects:
##      Groups   Name                Variance Std.Dev. Corr
##      subj      (Intercept)  0.0778564  0.27903
##      cond      cond          0.0018203  0.04267  0.28
##      item      condxxgroup  0.0003274  0.01809
##      item.1    cond          0.0001649  0.01284
##      item.2    group         0.0002864  0.01692
##      item.3    (Intercept)  0.0012332  0.03512
##      Residual                0.0935283  0.30582
## Number of obs: 2837, groups:  subj, 71; item, 20
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)    7.2811717  0.0450480 161.631
## cond           0.0637363  0.0105026   6.069
## group          0.3630934  0.0445195   8.156
## condxxgroup     0.0002732  0.0108825   0.025
## c_age          0.0033661  0.0029604   1.137
## cond:c_age     -0.0005562  0.0006742  -0.825
## group:c_age    -0.0091770  0.0029604  -3.100
## condxxgroup:c_age -0.0014003  0.0006742  -2.077
##
## Correlation of Fixed Effects:
##              (Intr) cond   group  cndxxg c_age  cnd:c_ grp:c_
## cond          0.176
## group         0.567  0.103

```

```
## condxxgroup  0.098  0.516  0.171
## c_age        -0.479 -0.087 -0.544 -0.094
## cond:c_age   -0.089 -0.468 -0.101 -0.507  0.185
## group:c_age  -0.538 -0.097 -0.485 -0.084  0.756  0.140
## cndxxgrp:c_  -0.100 -0.526 -0.090 -0.451  0.140  0.756  0.185
```