# Improving memory for and production of singular they pronouns: Experiment 1B

Bethany Gardner

03/24/2022

## Contents

This is a replication of the first experiment. Everything is identical, except that the production task occurs before the memory task.

## Load Data

Read data, preprocessed from Qualtrics output. See data/exp1b_data_readme for more details.

```
d <- read.csv("../data/exp1b_data.csv", stringsAsFactors=TRUE)
str(d)
```

```
## 'data.frame':    5151 obs. of  17 variables:
##  $ SubjID    : Factor w/ 101 levels "R_10uYBrhUOzbUAGV",..: 82 82 82 82 82 82 82 82 82 82 ...
##  $ SubjAge   : int  18 18 18 18 18 18 18 18 18 18 ...
##  $ SubjGender : Factor w/ 14 levels "female","Female",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ SubjEnglish: Factor w/ 3 levels "Fully competent in speaking listening reading and writing but not
##  $ List      : int  2 2 2 2 2 2 2 2 2 2 ...
##  $ Task      : Factor w/ 3 levels "introduction",..: 2 2 2 3 2 2 2 3 2 2 ...
```

```
##  $ Name       : Factor w/ 12 levels "Amanda","Andrew",..: 3 3 3 3 8 8 8 8 4 4 ...
##  $ Pronoun    : Factor w/ 3 levels "he/him","she/her",..: 1 1 1 1 1 1 1 1 3 3 ...
##  $ Pet        : Factor w/ 3 levels "cat","dog","fish": 1 1 1 1 2 2 2 2 3 3 ...
##  $ Job        : Factor w/ 12 levels "accountant","doctor",..: 12 12 12 12 4 4 4 4 9 9 ...
##  $ M_Type     : Factor w/ 4 levels "","job","pet",..: 2 3 4 1 2 3 4 1 2 3 ...
##  $ M_Response : Factor w/ 19 levels "","accountant",..: 19 7 9 1 8 5 9 1 14 7 ...
##  $ M_Acc      : int  0 0 1 NA 1 1 1 NA 1 1 ...
##  $ P_Response : Factor w/ 755 levels "","After Amanda got home from working as an engineer they took
##  $ P_Pronoun  : Factor w/ 5 levels "","he/him","none",..: 1 1 1 2 1 1 1 5 1 1 ...
##  $ P_Acc      : int  NA NA NA 1 NA NA NA 0 NA NA ...
##  $ I_Response : Factor w/ 297 levels "","Amanda goes by she/her. Amanda is a salesman and has a pet
```

Set up contrast coding. The first contrast compares they to he+she. The second contrast compares he to she.

```
contrasts(d$Pronoun) = cbind("they vs he+she"=c(.33,.33,-.66),
                             "he vs she"=c(-.5,.5, 0))
contrasts(d$Pronoun)
```

```
##            they vs he+she he vs she
## he/him               0.33      -0.5
## she/her              0.33       0.5
## they/them           -0.66       0.0
```

Split data by task, and only keep pronoun questions (not the job or pet questions) in memory dataframe.

```
m <- d %>% filter(M_Type=="pronoun")
p <- d %>% filter(Task=="production")
```

Combine memory and production trials to make one row for each character.

```
#Get pronoun memory and production observations. Filter out memory for job and pet questions, and intro
mp <- d %>% filter(Task != "introduction" &
                   M_Type != "job" &
                   M_Type != "pet")

#Just take columns used in model
m_temp <- mp %>% select(M_Acc, Pronoun, Name, SubjID) %>%
  filter(!is.na(M_Acc)) #Take out empty rows that were other question types

#Get production accuracy column
p_temp <- mp %>% select(P_Acc)  %>%
    filter(!is.na(P_Acc)) #Take out empty rows that were other question types

#Combine
mp <- cbind(m_temp, p_temp)
str(mp)
```

```
## 'data.frame':    1212 obs. of  5 variables:
##  $ M_Acc  : int  1 1 1 1 1 1 1 1 1 0 1 ...
##  $ Pronoun: Factor w/ 3 levels "he/him","she/her",..: 1 1 3 3 2 2 1 2 2 1 ...
##    ..- attr(*, "contrasts")= num [1:3, 1:2] 0.33 0.33 -0.66 -0.5 0.5 0
```

```
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : chr [1:3] "he/him" "she/her" "they/them"
##   .. .. ..$ : chr [1:2] "they vs he+she" "he vs she"
##  $ Name   : Factor w/ 12 levels "Amanda","Andrew",..: 3 8 4 9 10 12 2 5 1 11 ...
##  $ SubjID : Factor w/ 101 levels "R_10uYBrhUOzbUAGV",..: 82 82 82 82 82 82 82 82 82 82 ...
##  $ P_Acc  : int  1 0 1 1 1 1 1 1 1 1 ...
```

# Memory

**Descriptive Stats**

Mean accuracy for all three memory question types.

```
prop.table(table(d$M_Type, d$M_Acc), margin=1)
```

```
##
##                 0         1
##
##   job     0.7079208 0.2920792
##   pet     0.5717822 0.4282178
##   pronoun 0.3094059 0.6905941
```

Mean accuracy, split by pronoun type.

```
prop.table(table(m$Pronoun, m$M_Acc), margin=1)
```

```
##
##                   0         1
##   he/him    0.2128713 0.7871287
##   she/her   0.1980198 0.8019802
##   they/them 0.5173267 0.4826733
```

96% of participants selected they/them at least once.

```
they_m <- d %>% filter(M_Response=="they/them") %>%
  summarize(n=n_distinct(SubjID))

they_m/(n_distinct(d$SubjID))
```

```
##          n
## 1 0.960396
```

**Model**

Start with model that has random intercepts and slopes for participant and item. Maximal model has
by-participant random intercepts only.

```r
model_m_full <- M_Acc ~ Pronoun + (1 + Pronoun|SubjID) + (1 + Pronoun|Name)

model_m <- buildmer(model_m_full, data=m,
                    family='binomial', direction=c('order'))
```

## Determining predictor order

## Fitting via glm: M_Acc ~ 1

## Currently evaluating LRT for: Pronoun

## Fitting via glm: M_Acc ~ 1 + Pronoun

## Updating formula: M_Acc ~ 1 + Pronoun

## Fitting via glm: M_Acc ~ 1 + Pronoun

## Currently evaluating LRT for: 1 | Name, 1 | SubjID

## Fitting via glmer, with ML: M_Acc ~ 1 + Pronoun + (1 | Name)

## boundary (singular) fit: see help('isSingular')

## Fitting via glmer, with ML: M_Acc ~ 1 + Pronoun + (1 | SubjID)

## Updating formula: M_Acc ~ 1 + Pronoun + (1 | SubjID)

## Currently evaluating LRT for: 1 | Name, Pronoun | SubjID

## Fitting via glmer, with ML: M_Acc ~ 1 + Pronoun + (1 | SubjID) + (1 |
##      Name)

## boundary (singular) fit: see help('isSingular')

## Fitting via glmer, with ML: M_Acc ~ 1 + Pronoun + (1 + Pronoun |
##      SubjID)

## Ending the ordering procedure due to having reached the maximal
##      feasible model - all higher models failed to converge. The types of
##      convergence failure are: Singular fit lme4 reports not having
##      converged (-1)

```r
summary(model_m)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) (p-values based on Wald z-scores) [glmerMod]
##  Family: binomial  ( logit )
## Formula: M_Acc ~ 1 + Pronoun + (1 | SubjID)
##     Data: m
##
##      AIC      BIC   logLik deviance df.resid
##   1368.6   1389.0   -680.3   1360.6     1208
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.8110 -0.8346  0.4172  0.5693  1.7603
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  SubjID (Intercept) 0.3804   0.6168
## Number of obs: 1212, groups:  SubjID, 101
##
## Fixed effects:
##                      Estimate Std. Error  z value Pr(>|z|) Pr(>|t|)
## (Intercept)           0.94930    0.09466 10.02807    0.000  <2e-16 ***
## Pronounthey vs he+she 1.55205    0.14263 10.88181    0.000  <2e-16 ***
## Pronounhe vs she      0.09637    0.17857  0.53971    0.589    0.589
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) Prnvh+
## Prnnthyvsh+ 0.181
## Prononhvssh 0.018  0.018
```

Convert to odds:

```
exp(0.94930) #intercept (mean)
```

```
## [1] 2.5839
```

```
exp(1.55205) #they/them vs. he/him + she/her
```

```
## [1] 4.721139
```

- The intercept is significant (p<.001), such that participants are 2.58 times more likely to answer correctly than incorrectly across all pronoun types.

- The contrast between they/them and he/him + she/her is significant (p<.001), such that participants are 4.72 times more likely to get he/him and she/her right than they/them.

- The contrast between he/him and she/her is not significant.

# Production

**Descriptive Stats**

Mean accuracy, split by pronoun type. Accuracy for producing they/them is lower than accuracy for remembering they/them.

```
prop.table(table(p$Pronoun, p$P_Acc), margin=1)
```

```
##
##                        0          1
##   he/him     0.1584158 0.8415842
##   she/her    0.1559406 0.8440594
##   they/them  0.6089109 0.3910891
```

71% of participants produced they/them at least once.

```
they_p <- d %>% filter(P_Pronoun=="they/them") %>%
  summarize(n=n_distinct(SubjID))

they_p/(n_distinct(d$SubjID))
```

```
##           n
## 1 0.7128713
```

**Model**

Start with model that has random intercepts and slopes for participant and item, using same specifications as before. Maximal model has random intercepts by participant and item, and no random slopes.

```
model_p_full <- P_Acc ~ Pronoun +
  (1 + Pronoun|SubjID) + (1 + Pronoun|Name)

model_p <- buildmer(model_p_full, data=p,
          family='binomial', direction=c('order'))
```

```
## Determining predictor order

## Fitting via glm: P_Acc ~ 1

## Currently evaluating LRT for: Pronoun

## Fitting via glm: P_Acc ~ 1 + Pronoun

## Updating formula: P_Acc ~ 1 + Pronoun

## Fitting via glm: P_Acc ~ 1 + Pronoun

## Currently evaluating LRT for: 1 | Name, 1 | SubjID
```

```
## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + (1 | Name)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + (1 | SubjID)
```

```
## Updating formula: P_Acc ~ 1 + Pronoun + (1 | SubjID)
```

```
## Currently evaluating LRT for: 1 | Name, Pronoun | SubjID
```

```
## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + (1 | SubjID) + (1 |
##       Name)
```

```
## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + (1 + Pronoun |
##       SubjID)
```

```
## Updating formula: P_Acc ~ 1 + Pronoun + (1 | SubjID) + (1 | Name)
```

```
## Currently evaluating LRT for: Pronoun | Name, Pronoun | SubjID
```

```
## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + (1 | SubjID) + (1 +
##       Pronoun | Name)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + (1 + Pronoun |
##       SubjID) + (1 | Name)
```

```
## Ending the ordering procedure due to having reached the maximal
##       feasible model - all higher models failed to converge. The types of
##       convergence failure are: Singular fit lme4 reports not having
##       converged (-1)
```

```
summary(model_p)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##    Approximation) (p-values based on Wald z-scores) [glmerMod]
##  Family: binomial  ( logit )
## Formula: P_Acc ~ 1 + Pronoun + (1 | SubjID) + (1 | Name)
##     Data: p
##
##      AIC      BIC   logLik deviance df.resid
##   1207.4   1232.9   -598.7   1197.4     1207
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.0720 -0.6925  0.3443  0.4193  2.1954
##
## Random effects:
##  Groups Name         Variance Std.Dev.
```

```
##  SubjID (Intercept) 0.77891  0.88256
##  Name   (Intercept) 0.00775  0.08803
## Number of obs: 1212, groups:  SubjID, 101; Name, 12
##
## Fixed effects:
##                      Estimate Std. Error   z value Pr(>|z|) Pr(>|t|)
## (Intercept)          1.108332   0.123498  8.974463    0.000   <2e-16 ***
## Pronounthey vs he+she 2.476427  0.167460 14.788192    0.000   <2e-16 ***
## Pronounhe vs she      0.002089  0.223504  0.009344    0.993    0.993
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) Prnvh+
## Prnnthyvsh+  0.246
## Prononhvssh -0.025 -0.042
```

Convert to odds:

```
exp(1.10833) #intercept (mean)
```

```
## [1] 3.029295
```

```
exp(2.476427) #they/them vs. he/him + she/her
```

```
## [1] 11.89867
```

- The intercept is significant (p<.001), such that participants are 3.02 times more likely to answer correctly than incorrectly across all pronoun types.

- The contrast between they/them and he/him + she/her is significant (p<.001), such that participants are 11.83 times more likely to get he/him and she/her right than they/them.

- The contrast between he/him and she/her is not significant.

## Memory Predicting Production

**Descriptive Stats**

Combining the two measures, there are 4 possible patterns: getting both right, getting both wrong, getting just memory right, and getting just production right.

```
mp_acc <- mp %>%
        mutate(BothRight=ifelse(M_Acc==1 & P_Acc==1, 1, 0)) %>%
        mutate(BothWrong=ifelse(M_Acc==0 & P_Acc==0, 1, 0)) %>%
        mutate(MemOnly=ifelse(M_Acc==1 & P_Acc==0, 1, 0)) %>%
        mutate(ProdOnly=ifelse(M_Acc==0 & P_Acc==1, 1, 0)) %>%
        pivot_longer(cols=c(BothRight, BothWrong, MemOnly, ProdOnly),
                     names_to="Combined_Accuracy") %>%
        group_by(Pronoun, Combined_Accuracy) %>%
        summarise(m=mean(value))
```

```
## 'summarise()' has grouped output by 'Pronoun'. You can override using the
## '.groups' argument.
```

```
mp_acc
```

```
## # A tibble: 12 x 3
## # Groups:   Pronoun [3]
##    Pronoun   Combined_Accuracy      m
##    <fct>     <chr>               <dbl>
##  1 he/him    BothRight           0.688
##  2 he/him    BothWrong           0.0594
##  3 he/him    MemOnly             0.0990
##  4 he/him    ProdOnly            0.153
##  5 she/her   BothRight           0.708
##  6 she/her   BothWrong           0.0619
##  7 she/her   MemOnly             0.0941
##  8 she/her   ProdOnly            0.136
##  9 they/them BothRight           0.309
## 10 they/them BothWrong           0.436
## 11 they/them MemOnly             0.173
## 12 they/them ProdOnly            0.0817
```

**Model**

Model predicting production accuracy with pronoun type and memory accuracy. Here, the maximal model includes no random effects.

```
model_mp_full <- P_Acc ~ M_Acc * Pronoun +
  (1 + Pronoun|SubjID) + (1 + Pronoun|Name)

model_mp <- buildmer(model_mp_full, data=mp,
            family='binomial', direction=c('order'))
```

```
## Determining predictor order
```

```
## Fitting via glm: P_Acc ~ 1
```

```
## Currently evaluating LRT for: M_Acc, Pronoun
```

```
## Fitting via glm: P_Acc ~ 1 + M_Acc
```

```
## Fitting via glm: P_Acc ~ 1 + Pronoun
```

```
## Updating formula: P_Acc ~ 1 + Pronoun
```

```
## Currently evaluating LRT for: M_Acc
```

```
## Fitting via glm: P_Acc ~ 1 + Pronoun + M_Acc
```

```
## Updating formula: P_Acc ~ 1 + Pronoun + M_Acc
```

```
## Currently evaluating LRT for: M_Acc:Pronoun

## Fitting via glm: P_Acc ~ 1 + Pronoun + M_Acc + M_Acc:Pronoun

## Updating formula: P_Acc ~ 1 + Pronoun + M_Acc + M_Acc:Pronoun

## Fitting via glm: P_Acc ~ 1 + Pronoun + M_Acc + M_Acc:Pronoun

## Currently evaluating LRT for: 1 | Name, 1 | SubjID

## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + M_Acc + Pronoun:M_Acc
##      + (1 | Name)

## boundary (singular) fit: see help('isSingular')

## Fitting via glmer, with ML: P_Acc ~ 1 + Pronoun + M_Acc + Pronoun:M_Acc
##      + (1 | SubjID)

## Ending the ordering procedure due to having reached the maximal
##      feasible model - all higher models failed to converge. The types of
##      convergence failure are: Singular fit lme4 reports not having
##      converged (-1)
```

```
summary(model_mp)
```

```
##
## Call:
## stats::glm(formula = P_Acc ~ 1 + Pronoun + M_Acc + M_Acc:Pronoun,
##     family = "binomial", data = mp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0703  -0.5863   0.4995   0.5185   1.9214
##
## Coefficients:
##                           Estimate Std. Error z value Pr(>|z|)
## (Intercept)                0.02119    0.12994   0.163 0.870477
## Pronounthey vs he+she      2.56843    0.25749   9.975  < 2e-16 ***
## Pronounhe vs she          -0.16062    0.34055  -0.472 0.637175
## M_Acc                      1.49113    0.16078   9.274  < 2e-16 ***
## Pronounthey vs he+she:M_Acc -1.15555  0.32239  -3.584 0.000338 ***
## Pronounhe vs she:M_Acc     0.24029    0.41760   0.575 0.565016
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1496.3  on 1211  degrees of freedom
## Residual deviance: 1113.0  on 1206  degrees of freedom
## AIC: 1125
##
## Number of Fisher Scoring iterations: 4
```

Convert to odds:

```r
exp(1.49113)  #memory accuracy
```

```
## [1] 4.442112
```

```r
exp(-1.15555) #they/them vs. he/him + she/her * memory accuracy
```

```
## [1] 0.3148843
```

- The effect of memory accuracy is significant (p<.001), such that participants are 4.44x more likely to get the production right if they got the memory right.

- Significant interaction between pronoun type (they/them vs. he/him + she/her) and memory accuracy (p<.05) (odds 0.31). The relative difficulty of they/them was attenuated when the participant had correctly remembered the character's pronoun during the memory phase of the task.