Physical Pragmatics (Experiment 2c)

Preprocessing

```
# Read in the participant data.
data_0 = read_csv(file.path(data_path, "raw_data.csv"), quote="~")

# Convert the JSON string into JSON.
data_1 = lapply(data_0$data, fromJSON)

# Extract the trial information for each participant and stack them.
data_3 = tibble()
for (p in 1:length(data_1)) {
  # Trim the map and add the participant ID back in.
  data_2 = data_1[p][[1]]$trials %>%
    as.data.frame() %>%
    separate(stimuli, into=c("door", "cost", "object"), sep="_") %>%
    mutate(object=gsub(".png", "", object),
           target=as.numeric(target),
           unique_id=data_1[p][[1]]$id,
           age=as.numeric(data_1[p][[1]]$subject_information$age))

  # Stack the trial information for the current participant.
  data_3 = rbind(data_3, data_2)
}

# Write the preprocessed data.
write_csv(data_3, file.path(data_path, "data.csv"))
```

"Unusualness" as a predictor of communicative meaning

Now, we analyze the predictive power of participant "unusualness" judgments (N=80, M=34.04 years, SD=15.1 years) on the participant judgments from Experiment 2a. This time, we'll also include the cost condition as a predictor. If both predictors are significant, we will run a second analysis to determine whether or not the cost condition can explain the unexplained variance from participant "unusualness" judgments.

```
# Read in the preprocessed participant data.
data_3 = read_csv(file.path(data_path, "data.csv"))

# Compute the mean participant judgments.
data_4 = data_3 %>%
  select(cost, object, target) %>%
  group_by(cost, object) %>%
  summarize(unusualness=mean(target))
```

```r
# Read in the preprocessed data for the first partition.
data_5 = read_csv("../experiment_2a/data/data_0/data.csv") %>%
  rbind(read_csv("../experiment_2a/data/data_1/data.csv"))

# Exclude participants who said the unmodified door was more difficult to walk
# through and chop off the extra participant in the low-cost condition.
data_6 = data_5 %>%
  filter(costlier!="unmodified", participant!=161)

# Filter the columns of interest and append the unusualness judgments.
data_7 = data_6 %>%
  rename(cost=condition) %>%
  select(response, cost, object) %>%
  full_join(data_4)

# Compute a logistic regression predicting `decider_0` participant judgments as
# a function of participant "unusualness" judgments and the cost condition.
model_0 = glm(response~unusualness+cost, data=data_7, family="binomial")
summary(model_0)
```

```
##
## Call:
## glm(formula = response ~ unusualness + cost, family = "binomial",
##     data = data_7)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3900  -1.0398   0.4870   0.7676   1.3216
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.4403     0.7027   0.627   0.5309
## unusualness   2.3756     1.0176   2.334   0.0196 *
## costnone     -0.9702     0.4707  -2.061   0.0393 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 184.21  on 159  degrees of freedom
## Residual deviance: 162.16  on 157  degrees of freedom
## AIC: 168.16
##
## Number of Fisher Scoring iterations: 4
```

Both predictors are significant, so we will run a series of regressions to test whether or not the cost condition can explain any of the unexplained variance left by "unusualness".

```r
# Compute a logistic regression predicting `decider_0` participant judgments as
# a function of participant "unusualness" judgments.
model_1 = glm(response~unusualness, data=data_7, family="binomial")
```

```r
# Extract the residuals (i.e., unexplained variance) from the previous model.
data_7$residuals = resid(model_1)

# Compute a linear regression predicting the residuals as a function of the
# cost condition.
model_2 = glm(residuals~cost, data=data_7, family="gaussian")
summary(model_2)
```

```
##
## Call:
## glm(formula = residuals ~ cost, family = "gaussian", data = data_7)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.6899  -0.9682   0.2654   0.6394   1.3953
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.3110     0.1120   2.776  0.00616 **
## costnone     -0.3619     0.1584  -2.284  0.02368 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 1.003894)
##
##     Null deviance: 163.85  on 159  degrees of freedom
## Residual deviance: 158.62  on 158  degrees of freedom
## AIC: 458.67
##
## Number of Fisher Scoring iterations: 2
```