acoustic_analysis

```
#
# location_condition is an attribute of the trials with levels Agent, patient, verb, and control.
# locaiton is an attribute of the word with levels Agent, patient, verb
# condition is an attribute of the trial with levels contrast and control.

cur_exp = "exp0"
features = c("duration", "meanIntensity", "meanpit")
# info = c('participant','verb','condition', 'word', 'word_num')
info = c('participant','item_id','location_condition', 'word', 'word_num', 'present_num')
bRemove_outliers = 0
# I have experimented with removing outliers, it doesn't have much effect on duration, some people with

tAll_trials = read.csv(file.path('..', cur_exp, 'tAll_trials.csv'))


# df0 = read.csv(paste0('measure_', cur_exp, '.csv'), header = T)
df0 = read.csv(paste0('measure_', cur_exp, '.csv'))

df0$location_condition = NA
df0$item_id = NA

for (iR in 1:nrow(df0)){
  df0$location_condition[iR] =  as.character(tAll_trials[tAll_trials$trial_id == df0$trialId[iR],'locat
  df0$item_id[iR] =  as.character(tAll_trials[tAll_trials$trial_id == df0$trialId[iR],'filler_or_item_i
  df0$present_num[iR] = as.numeric(rownames(tAll_trials[tAll_trials$trial_id == df0$trialId[iR],]))
}


df1 = df0[startsWith(df0$item_id, "item"),]

# df0 = read.csv("measure_nonrhyming_84total_60No_24Yes_20181210.csv", header  = T)
# df0 = transform(df0,trialId=as.numeric(trialId))
# sort(df0$trialId, decreasing = FALSE)
# colnamesC(df1)

df2 = df1[df1$word != 'sp',] # there can be sp everywhere not just begginning or end
# code for word_num
df2 <- df2 %>%
  dplyr::group_by(participant, trialId) %>%
  # dplyr::group_by(participant, question, trialId) %>%
  dplyr::mutate(word_num=1:dplyr::n()) %>%
  dplyr::select(c(info, features))
```

```
## Adding missing grouping variables: `trialId`
```

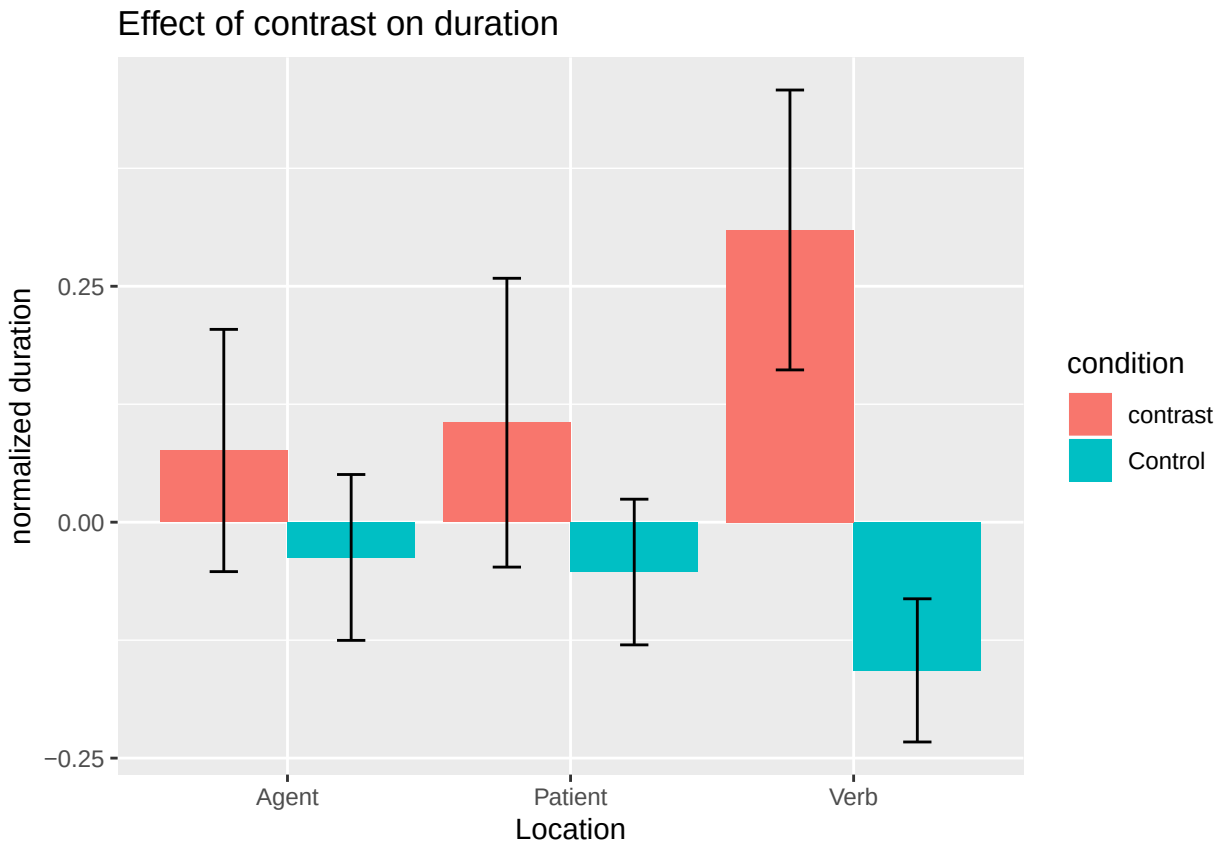
```
# c(df_Verb, df_Agent, df_Patient) %<-% process_data_with_yes(df2)
c(df_Verb, df_Agent, df_Patient) %<-% process_data_without_yes(df2)
```

```
## [1] 0
```

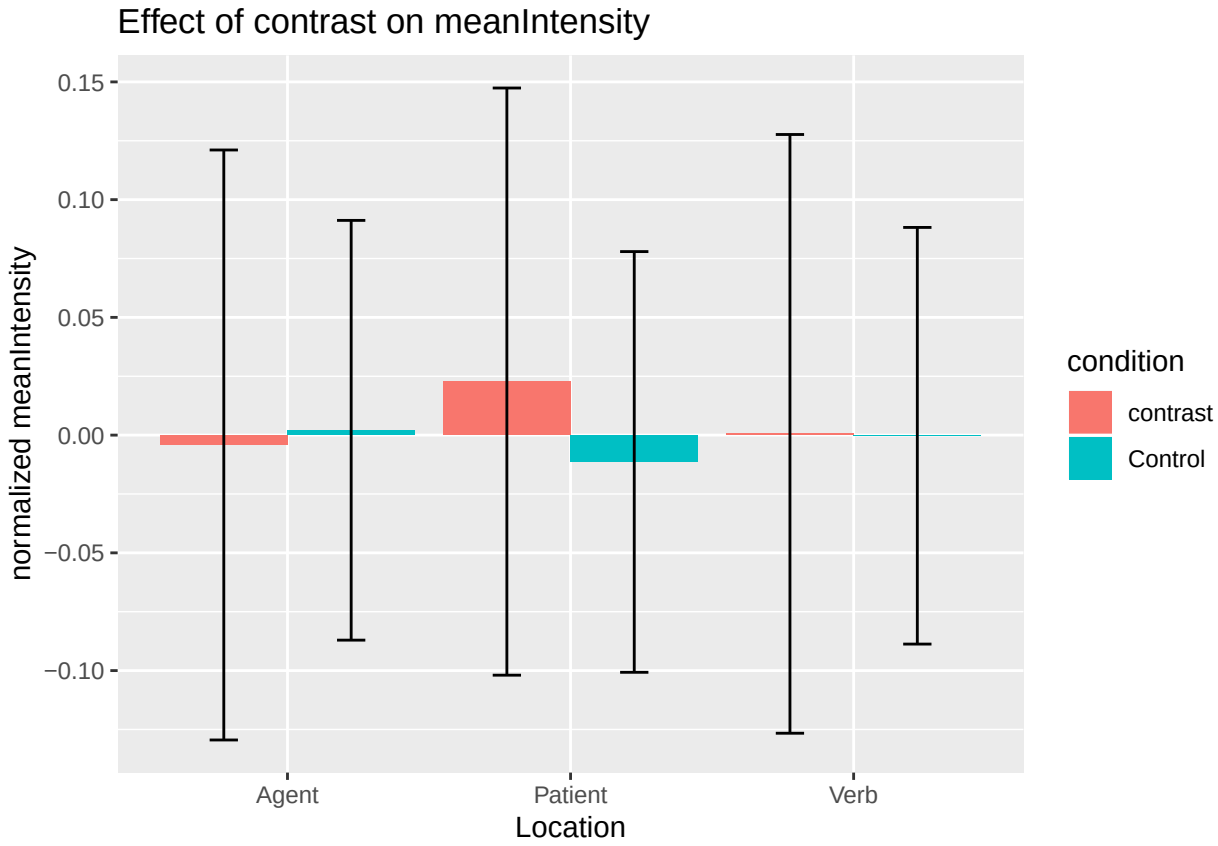
```
# the agent diff is smaller without yes than yes, maybe because when correcting, agent is stressed even
```

```
if(cur_exp %in% c("exp4", "exp6", "exp8")){
  df_Agent$condition = mapvalues(df_Agent$location_condition,c('Agent'),c('contrastive no'))
  df_Verb$condition = mapvalues(df_Verb$location_condition,c('Agent'),c('noncontrastive no'))
  df_Patient$condition = mapvalues(df_Patient$location_condition,c('Agent'),c('noncontrastive no'))
} else{
  df_Agent$condition = mapvalues(df_Agent$location_condition,c('Agent'),c('contrast'))
  df_Verb$condition = mapvalues(df_Verb$location_condition,c('Verb'),c('contrast'))
  df_Patient$condition = mapvalues(df_Patient$location_condition,c('Patient'),c('contrast'))
}


df_Agent$Location = 'Agent'
df_Verb$Location = 'Verb'
df_Patient$Location = "Patient"


combined_dataset = rbind(df_Agent,df_Verb,df_Patient)

# http://www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2)/

for (iF in features){
  print(iF)

  summarized_dataset = summarySE(combined_dataset,measurevar=iF ,groupvars=c('Location','condition'))

  print(
    ggplot(summarized_dataset, aes(x=Location, y=get(iF), fill=condition)) +
      geom_bar(position=position_dodge(), stat="identity") +
      geom_errorbar(aes(ymin=get(iF)-ci, ymax=get(iF)+ci),
                    width=.2,
                    position=position_dodge(.9))+
      xlab("Location") +
      ylab(paste0("normalized ", iF)) +
      # scale_fill_hue(name="location_condition",
      #                breaks=c("Control", "contrast"),
      #                labels=c("NonContrastive", "Contrastive")) +
      ggtitle(paste0('Effect of contrast on ', iF ))
  )
}
```

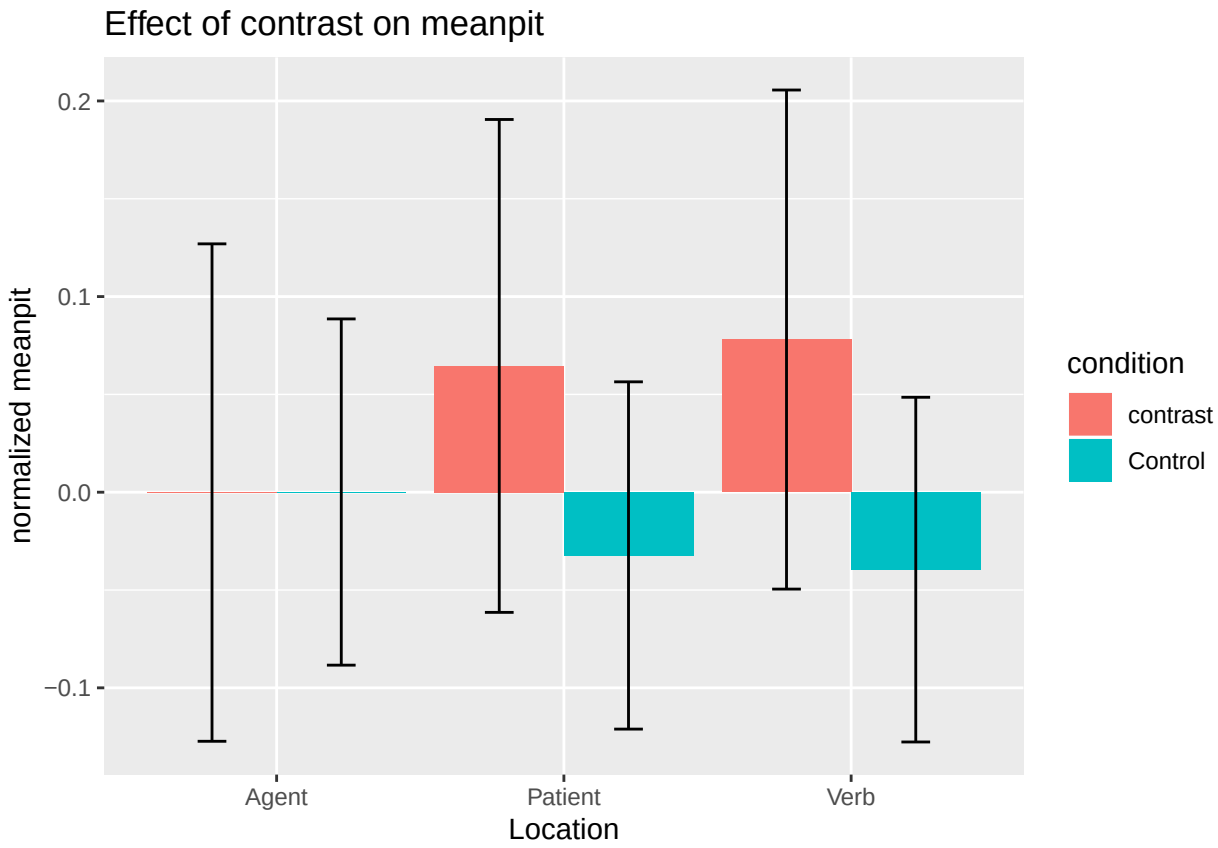
```
## [1] "duration"
```



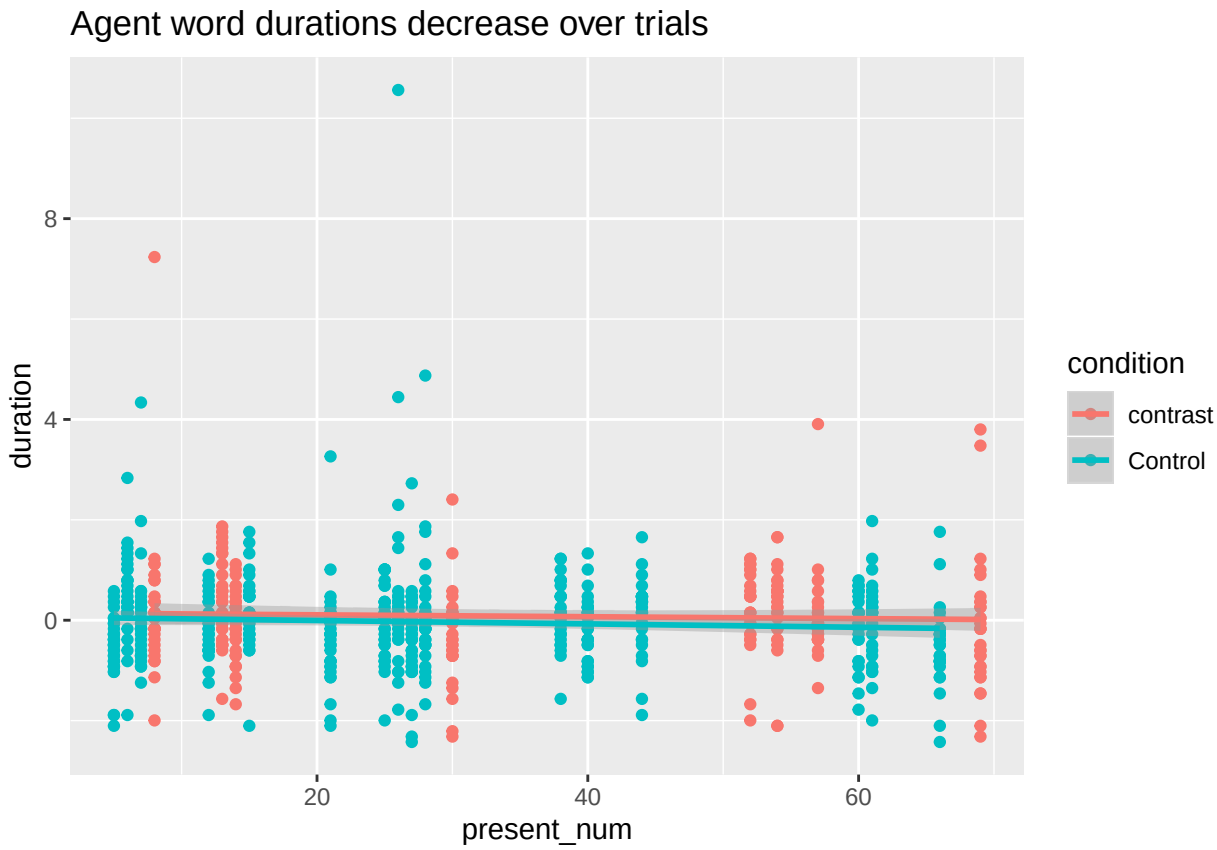

```
## [1] "meanIntensity"
```


Effect of contrast on meanIntensity
normalized meanIntensity
0.15
0.10
0.05
0.00
−0.05
−0.10
Agent
Patient
Verb
Location
condition
contrast
Control

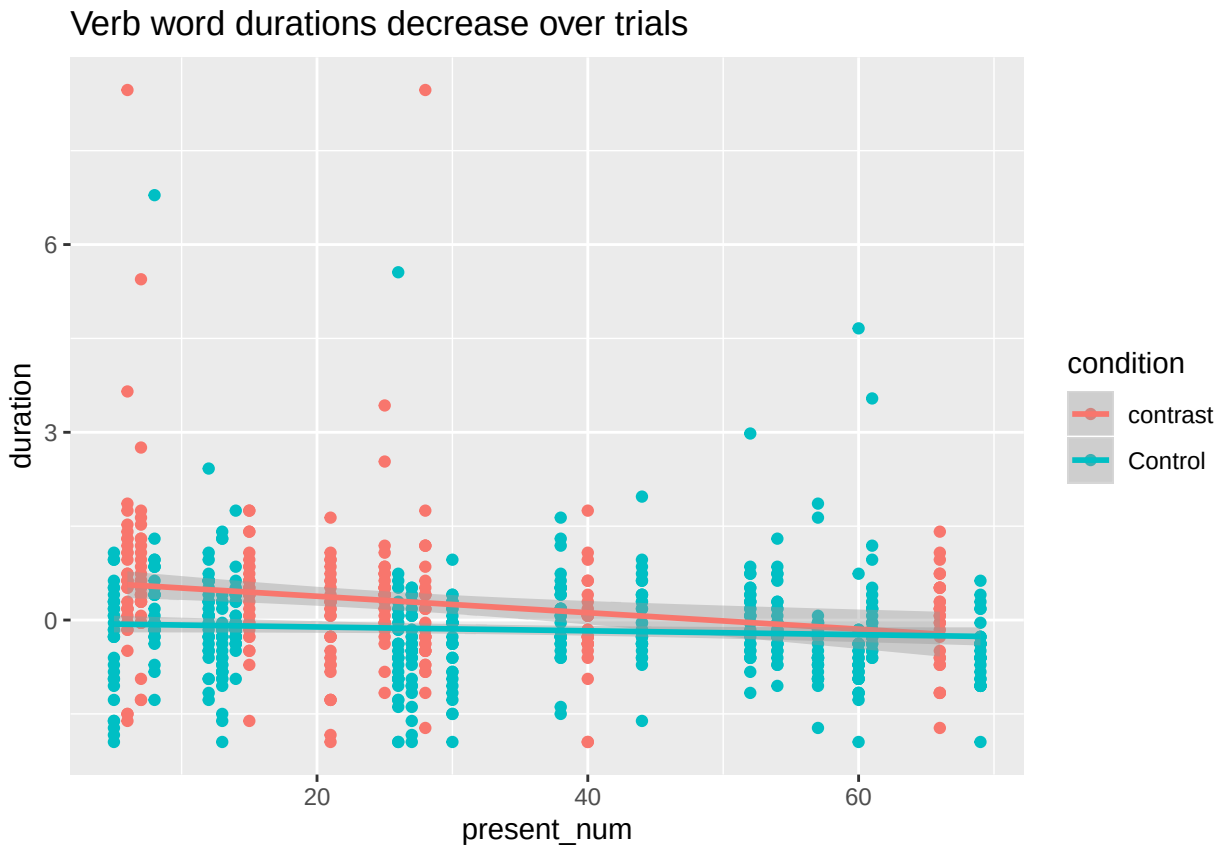
```
## [1] "meanpit"
```



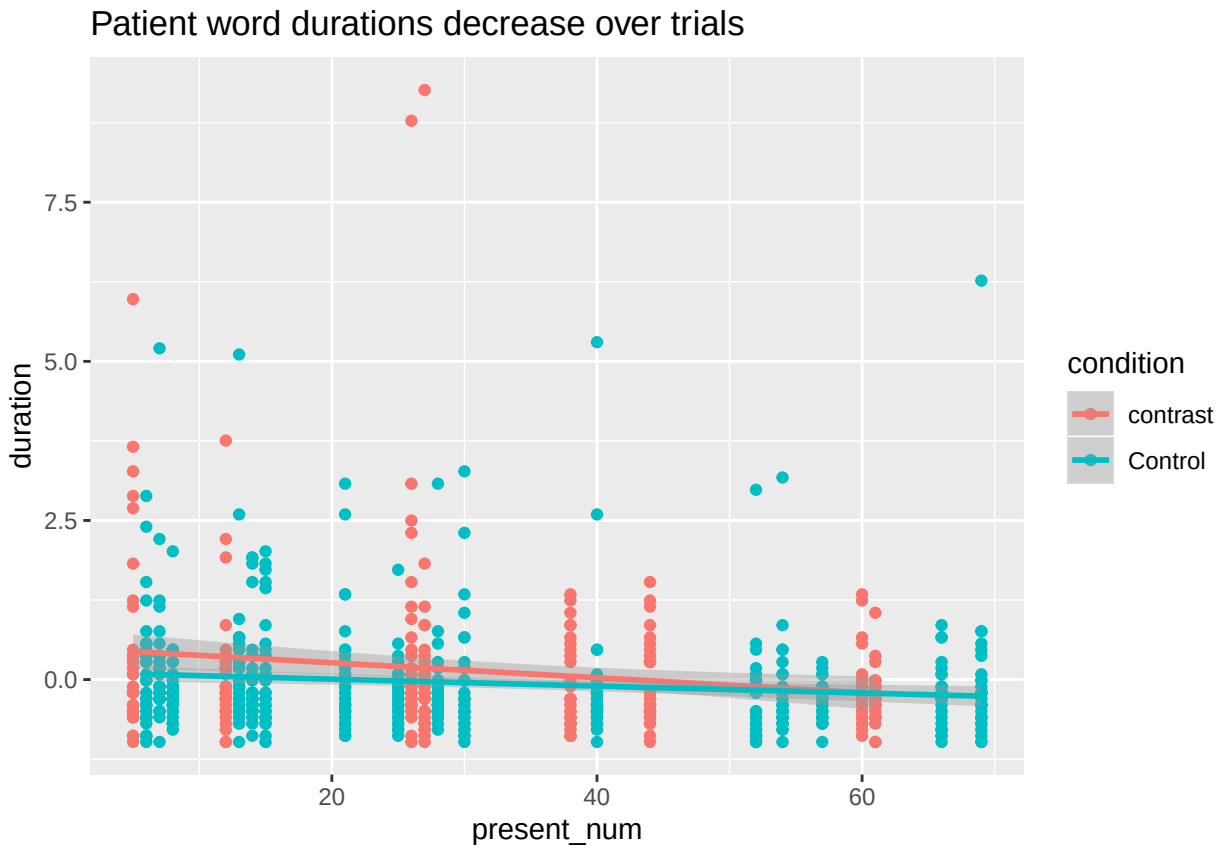

```
ggplot(df_Agent, aes(x=present_num, y=duration, color=condition)) +
  geom_point() + geom_smooth(method=lm) + ggtitle("Agent word durations decrease over trials")
```




```
ggplot(df_Verb, aes(x=present_num, y=duration, color=condition)) +
  geom_point() + geom_smooth(method=lm) + ggtitle("Verb word durations decrease over trials")
```




```
ggplot(df_Patient, aes(x=present_num, y=duration, color=condition)) +
  geom_point() + geom_smooth(method=lm) + ggtitle("Patient word durations decrease over trials")
```



32 workers and 1374.8 trials are included in this analysis.

This the analysis for exp0. The parameters of all exps can be seen at https://github.com/Xinzhu-Fang/prosody_study_exp/blob/master/tAll_exps.csv.

The trial-by-trial design of this exp can be seen at https://github.com/Xinzhu-Fang/prosody_study_exp/blob/master/exp0/tAll_trials.csv

Some code are hidden for the convenience of viewing results. Full code can be found at https://github.com/Xinzhu-Fang/prosody_study_exp/blob/master/analysis/acoustic_analysis.Rmd

```
# for (iF in features){ run_regression('Agent',iF)
# run_regression('Patient', iF) run_regression('Verb', iF) }
# r = lmer(get(observation) ~ condition + (1 | participant) +
# (1 | verb), data=df)
```