Two samples - two days

The word order in voice messages and text messages differs in the order of an object relative to a verb: in text messages the object is before the verb, in voice messages the verb is before the object. In languages with (relatively) free words order clauses with different words orders are observed. Nevertheless, even in such languages there is the basic or neutral words order. Words orders in these languages could be motivated by different information structure: for instance, a special position for the topic in the clause. However, informative structure is not the only reason for this. Formal vs informal communication or spoken vs written language may reveal distinctive structures. Moreover, we may expect variations in the word order of oral spoken colloquial language and written colloquial speech as its written representation. The aim of our research is to investigate it.

Data: voice messages and text messages annotated for the position of an object relative to a verb, texts similar in their topic and in size for each speaker (the thematic task for each respondent is the same).

H0: the basic word order (VO) does not depend on communicative strategy (in voice messages and text messages is the same). H1: the basic word order (VO) depends on communicative strategy (the word order in voice messages (VO) and in text messages (OV)).

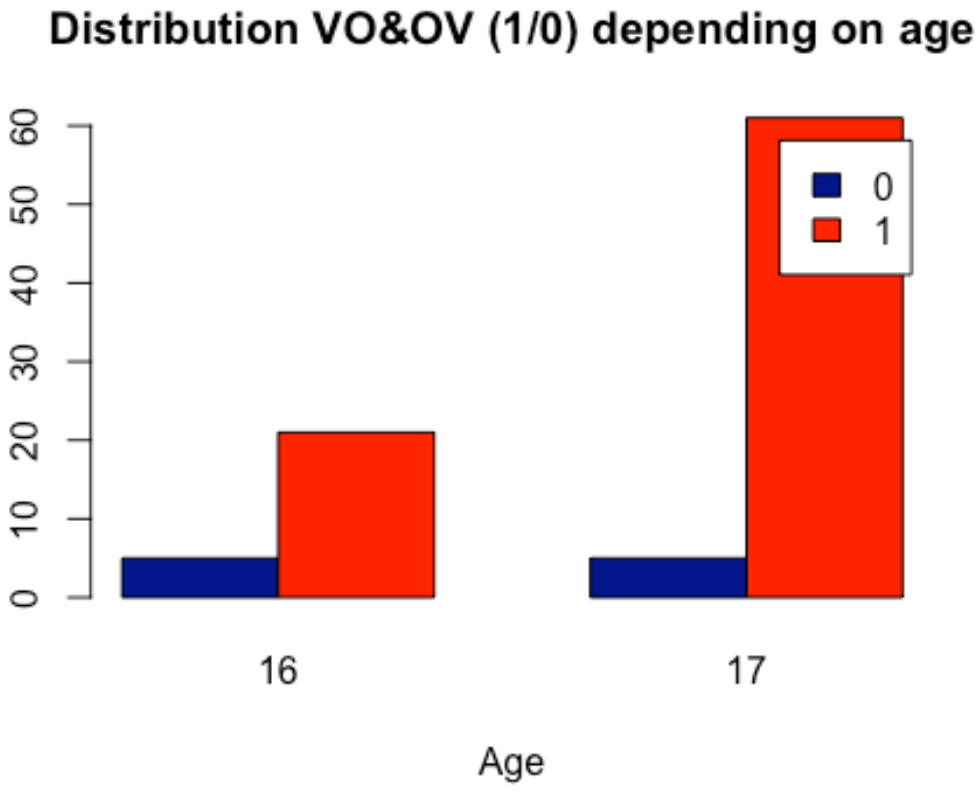
We asked students of the age of 15-18 years (students of HSE Lyceum of 10th and 11th grade) both to record a voice message and to write a text message on a messenger (Telegram) in a private chat. Consultants were divided into two groups: first group supposed to record a voice message and then to write a text message; the second group had to write a text message first and then to record a voice message. The topic was fixed: participants asked what they did yesterday. We also collected the following sociolinguistic information:

- gender
- age
- grade (10th/11th)
- written/spoken languages
- amount of messages
- the order of tasks
- city/town

The sample is not representative enough: the consultants seem to have no distinctive features. So, we decided to use the responses of 16 and 17 age consultants only.

After the data processing we discovered that the probability of basic word order potentially depends on age but not other factors that describe the consultants.

```
counts <- table(data$Target, data$Age)
barplot(counts, main="Distribution VO&OV (1/0) depending on age",
  xlab="Age", col=c("darkblue","red"),
  legend = rownames(counts), beside=TRUE)
```



According to this fact, we built the basic model:

```
model = glm('Target ~ Age', data=data[, c('Target', 'Age')], family =
'binomial')
summary(model)

##
## Call:
## glm(formula = "Target ~ Age", family = "binomial", data = data[,
##     c("Target", "Age")])
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2717   0.3969   0.3969   0.3969   0.6536
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.4351     0.4976   2.884  0.00393 **
## Age17         1.0664     0.6812   1.565  0.11748
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 63.255  on 91  degrees of freedom
## Residual deviance: 60.870  on 90  degrees of freedom
## AIC: 64.87
##
## Number of Fisher Scoring iterations: 5
```

Hence, the probability of basic word order does not significantly depend on age. However, one can see the potential tendency for more sofisticated samples.

```
model_message = glm('Target ~ Message.type * Age', data=data[, c('Target',
'Message.type', 'Age')], family = 'binomial')
summary(model_message)
```

```
##
## Call:
## glm(formula = "Target ~ Message.type * Age", family = "binomial",
##     data = data[, c("Target", "Message.type", "Age")])
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3678   0.3536   0.4366   0.4366   0.7876
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)           1.0116     0.5839   1.733   0.0832 .
## Message.typev         1.2910     1.2004   1.075   0.2822
## Age17                 1.7292     0.9344   1.851   0.0642 .
## Message.typev:Age17  -1.7292     1.5296  -1.130   0.2583
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 63.255  on 91  degrees of freedom
## Residual deviance: 59.295  on 88  degrees of freedom
## AIC: 67.295
##
## Number of Fisher Scoring iterations: 5
```

When adding the factor of communicative strategy type, the z-scores of random slope estimates and factor estimates itself are above significance levels. So, we can consider that communicative strategy does not influnce the word order of consultants and we have no reasons to reject H0.