

Project Proposal, 2020

Irina Astafyeva, Ksenia Romanova

A hypothesis.

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.

Research design.

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).

Formal hypotheses:

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))

Model: linear regression.

Statistical tools: Chi-squared test/Fisher exact test

Description of the data collection method.

We ask 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 social media/messenger. Consultants are divided into two groups: first group supposed to record a voice message and then to write a text message; the second group has to write a text message first and then to record a voice message. The topic is fixed: participants asked what they did yesterday. We will collect the following sociolinguistic information:

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