

# Prepositions in non-standard Russian in Finland: a learner corpus approach

---

Ekaterina Vlasova, [ekaterina.vlasova@helsinki.fi](mailto:ekaterina.vlasova@helsinki.fi)  
PhD student, University of Helsinki

# Outline

1. Non-standard Russian in Finland: teaching bilinguals and L2 Russian speakers
2. Russian prepositional phrase as an indicator of linguistic complexity
3. Method: combining corpus-based techniques and an error-focused approach
4. Prepositions as an indicator of the lexical diversity
5. Results
6. Pedagogical implications
7. Conclusion

# Non-standard Russian in Finland

Students with different backgrounds meet at the same Russian courses

- heritage Russian speakers from Russian speaking families (heritage learners - HL)
- advanced learners of L2 Russian (foreign learners - FL)

Advanced learners with unbalanced competence influenced by Finnish

How to improve effectively their linguistic competence acquired in different settings?

Descriptive and exploratory research of the learners' written production

# Non-standard Russian in Finland: the data

Russian learner corpus with manual annotation of errors

Finnish-Russian heritage corpus, 113, 027 tokens

Finnish-Russian L2 corpus, 201, 848 tokens

The data come from the University of Helsinki

The online repository <http://www.web-corpora.net/RLC/>

The project is a PhD dissertation supported by the Kone Foundation in 2018

# Preposition and lexical complexity

- 30 one-word prepositions from the core slavic lexicon
- 56 prepositions including multi-word expressions and adverbial derivatives  
(Sichinava 2018, Russian corpus grammar)

## Detailed conceptualization of space

- near *около - возле - вблизи - рядом с - недалеко от*
- through *сквозь - через*
- instead of *вместо - взамен*

# Preposition and morpho-syntactic complexity

In Russian, a preposition governs a grammatical case

**FOR**      dlja - Genitive

**TO**        k - Dative

**ABOUT**   o Prepositional

Some prepositions can govern a few cases:

**WITH** s and Instrumental

**FROM** s and Genitive

A learner of Russian is supposed to master about **30 - 70 combinations** of a preposition and its case

# Preposition and complex sentence

In Russian, a preposition is incorporated in a multi-word conjunction of a subordinate clause

Прежде чем... - before Ving

Перед тем как...- before Ving

После того как... - after Ving

Вместо того чтобы.... - instead of Ving

Для того чтобы... - for Ving

and so on

# Method

In my study, I use a Russian prepositional phrase as a test bed to explore how to

detect the areas of underacquisition (what else to learn)

detect compensatory overuse (what to correct)

evaluate the complexity of a linguistic phenomenon (when to learn)

evaluate the complexity of learner production (what has been learnt)

using the annotated learner data and corpus-based quantitative techniques



# Case study 1: Exploring lexical complexity

A list of preposition can be seen as a wordlist reflecting how many conceptual meanings a learner can express at a certain level. The longer the list of prepositions, the more detailed the conceptualization is.

By comparing frequency lists of prepositions, we can detect the gaps in lexical competence between heritage Russian speakers and advanced Russian l2 learners, as well as any other asymmetries and deviations from the standard Russian.

# The corpora to compare

FL - Russian L2 learners from Finland

HL - Russian heritage learners from Finland

KRUT - the corpus of Russian essays written by Russian university students

RNC - Russian National Corpus as a golden standard

# Quantitative techniques and procedure

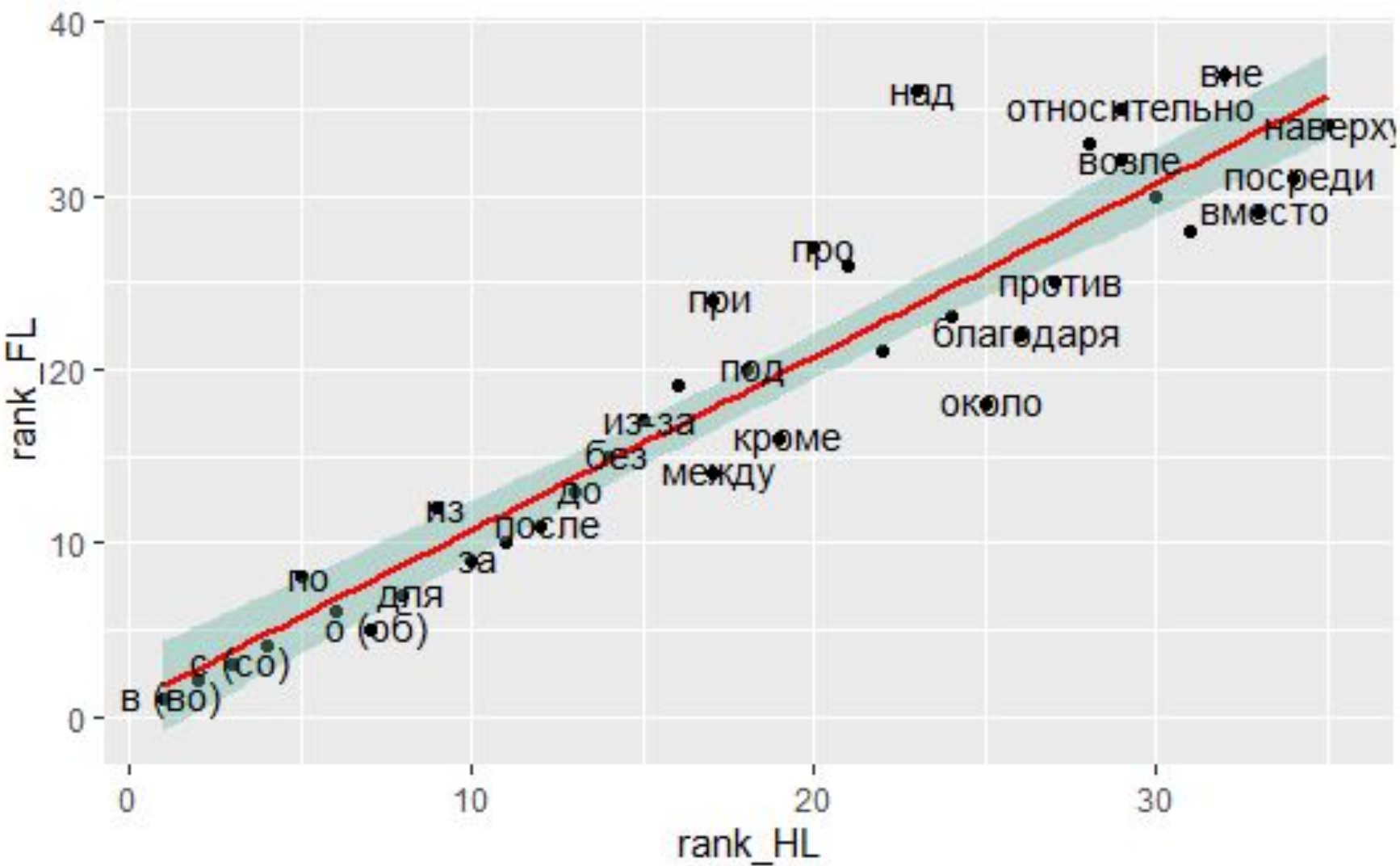
1. Compile a wordlist of prepositions for each corpora
2. Calculate normalized frequencies (items per million) and arrange the wordlist in descending order
3. Define the rank of the preposition in the frequency list according to the Zipfian rank frequency distribution: the higher the frequency, the lower the rank
4. Compile a table indicating the rank of the preposition in each corpus
5. Visualize the data using a scatterplot
6. Analyze the asymmetries and examine pedagogical implications
7. Build a linear regression model to verify the observed results statistically

## The frequency table with ranks

preposition	rank_FL	rank_RNC	rank_krut	rank_HL
без	15	16	16	14
благодаря	22	28	24	26
в (во)	1	1	1	1
вместо	29	25	32	33
вне	37	32	30	32
внутри	28	31	29	31
возле	32	26	36	29
вокруг	23	30	33	24
для	7	11	5	8

Asymmetries in the  
competence between L2  
and bilinguals

---



# Results

L2 learners underuse

при (concessive meaning)

про (spoken about)

над (above)

возле (near)

вне (out)

относительно (concerning)

L2 learners overuse

между between

кроме except

около near

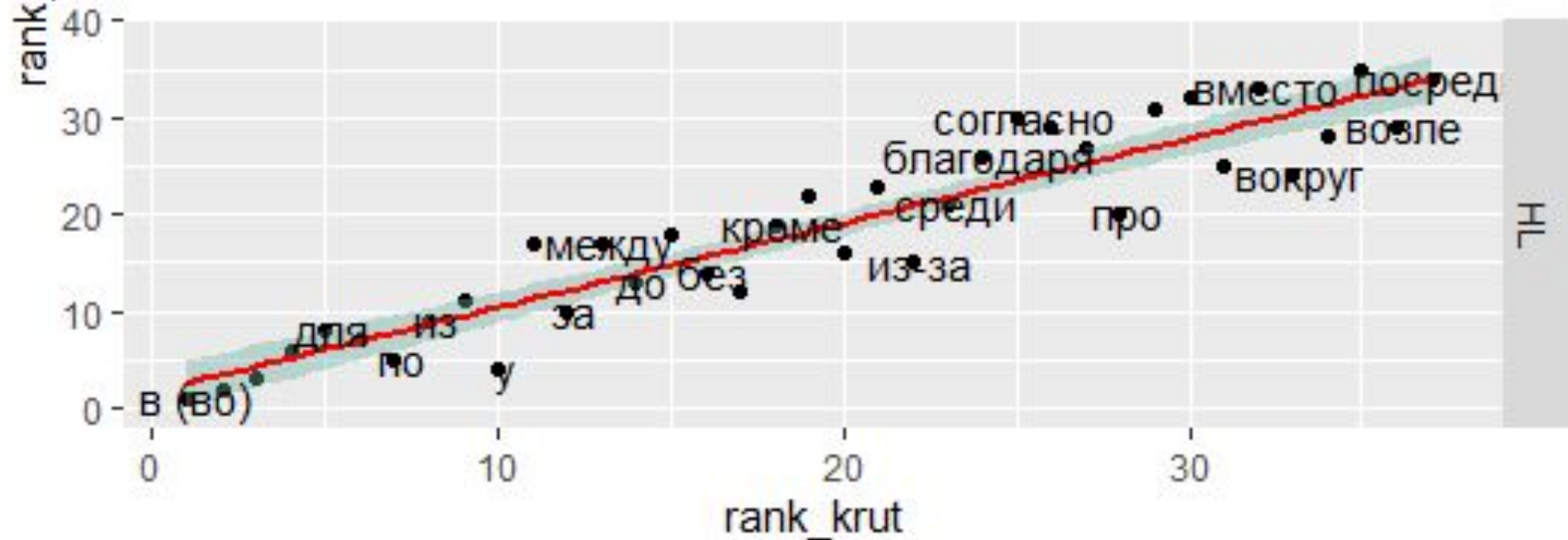
благодаря due to

вместо instead of

посреди midst

What are the asymmetries  
compared against standard  
Russian speakers?





# The results: L2 learners

L2 learners underuse

при (concessive meaning)

про (spoken about)

над (above)

возле (near)

вне (out)

относительно (concerning)

L2 speakers overuse

у

за

после

из-за

около

вокруг

# The results: heritage bilinguals

## **underuse**

при (concessive meaning)

между (between)

под (under)

перед (in front of)

над (above)

возле (near)

вне (out)

относительно (concerning)

благодаря

## **overuse**

по

у

за

после

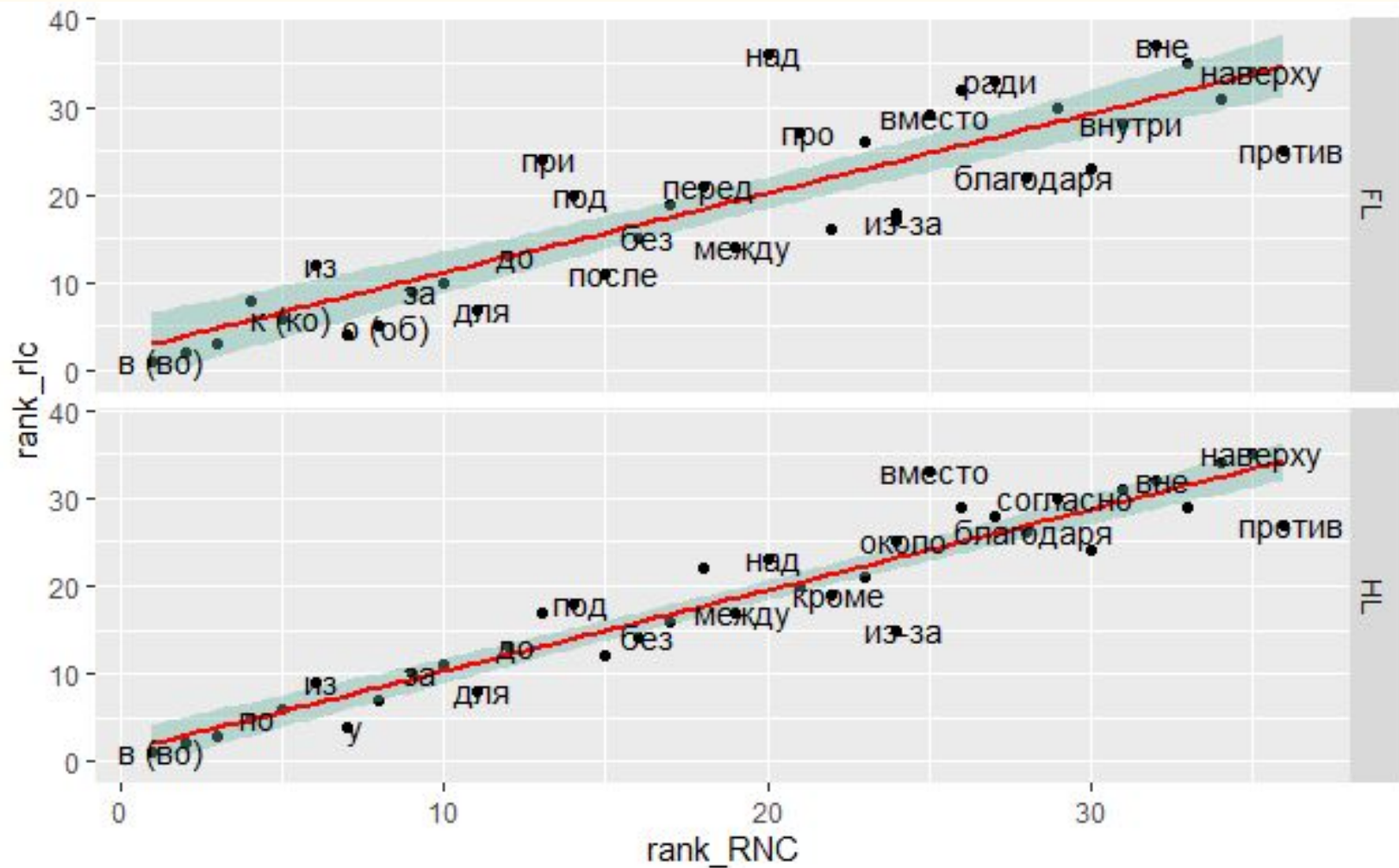
через

про

около

вокруг

What are the asymmetries  
compared against the  
Russian National Corpus?



# Overuse

L2 learners overuse

о(об)

для

после

между

heritage bilinguals overuse

у

для

из-за

# Underuse

L2 learners

при

под

над

возле

вместо

heritage speakers

между

перед

при

под

над

возле

вместо

## Examples of the overuse of по, для, о(heritage speakers)

Я часто работаю **по каникулам** летом и зимой.

Умный человек не сидит над компьютером **по судкам**.

Вот эти три последних пункта – дисциплина, укрепления здоровые и удовлетворение **причины для аргумента** что спорт полезен и подходит всем людям

но к счастью всем не нужно посвятить свою жизнь **для политики**.

Может быть хорошо, если в школе учились более **о политике**.



# Conclusion

The list of the prepositions and related constructions to teach the advanced level

- underuse of *при* , *относительно* - wide range of use
- prepositions denoting time and space: *перед* ‘in front of’ *над* ‘above’, *под* ‘below’
- correcting overuse of the prepositions *для* ‘for’, *о* ‘about’, *по* ‘by’ and others from the list.

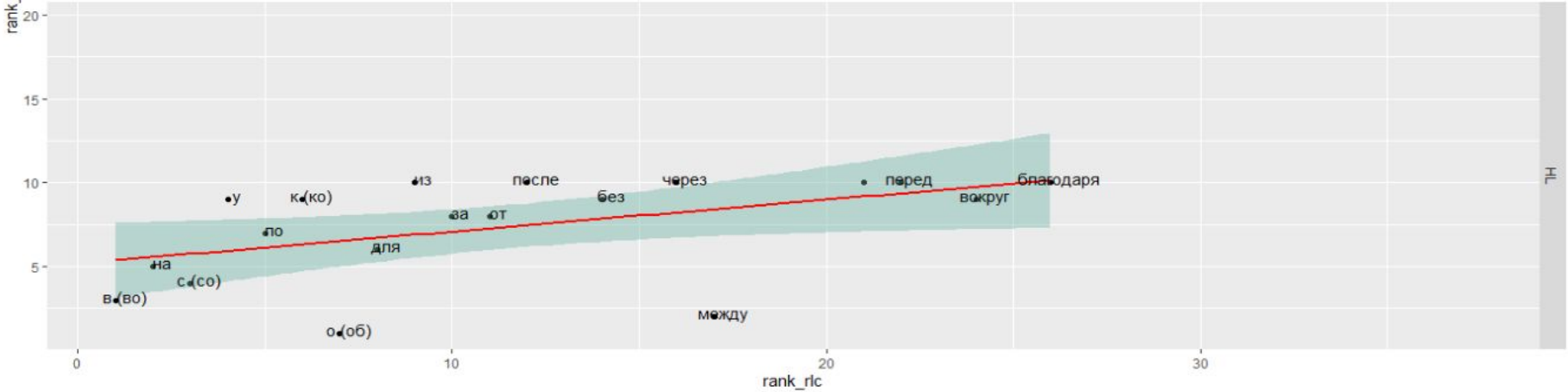
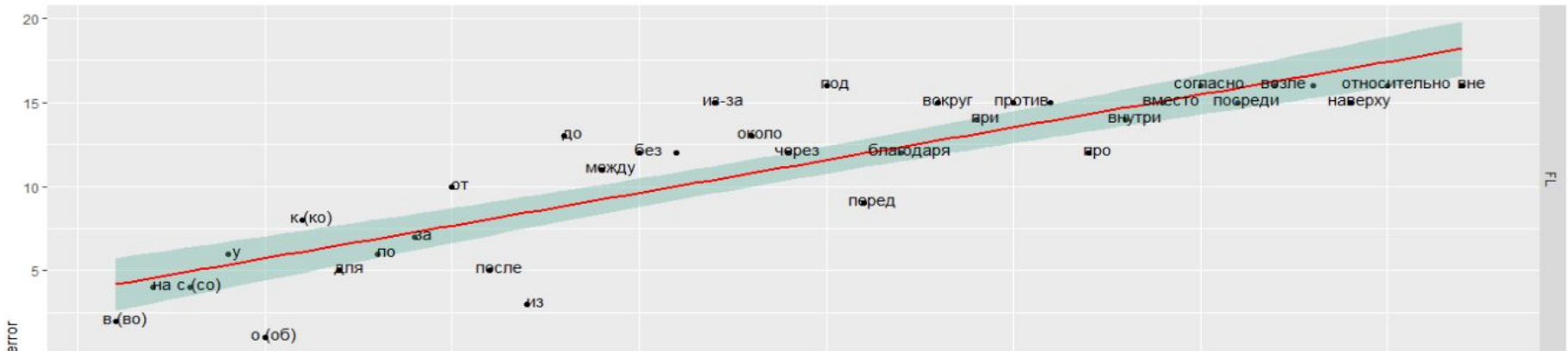
## Case study 2: grammatical simplification

Instead of 35-70 combinations of preposition and its case, Russian learners tend to simplify the case government within the prepositional phrase in a more unified way.

The most common morpho-syntactic error concerns overuse of Genitive with preposition, which normally demands other cases.

Is this process developing in the same way in bilingual and L2 production?

□ [38;5;232m'geom\_smooth() using formula = 'y ~ x'[39m



# Implications for teaching

- + easy technique for mapping out the usage on one slide
  - + seems to be effective for comparisons
  - + effective for detecting the lexical units with abnormal usage
- nuances are lost, and in-depth linguistic analysis is still needed to understand the nature of errors and generalizations underlying the observed overuse
- some asymmetries occur due to the variation of genre and topic rather than developmental process