

Annotation of Morphological Errors in L2 Russian Corpus Analysis

Irina KOR CHAHINE
Ekaterina UETOVA

University Côte d'Azur, Nice, France

contact: irina.kor-chahine@univ-cotedazur.fr

*21st Annual Second Language Acquisition and Teaching Interdisciplinary Roundtable,
Saturday, February 5th, 2022 – The University of Arizona, Tucson, AZ*

Introduction

- This study is a part of a research project *Grammar of Errors in Russian L2* devoted to analysis of students' errors in Russian L2.
- This is a cross-sectional study involving all language proficiency levels which is based on Russian L2 texts written by French students.
- The texts were collected, digitalized, and manually annotated in the Russian Learner Corpus (RLC, <http://www.web-corpora.net/RLC/>, free access)
- Then, the errors were analyzed quantitatively, as well as qualitatively.
- So, we will present here particularities of each stage of data processing.

Step 1: Data Collection

Our working corpus involves two corpora:

1. **Pilot corpus** with various L1 languages
(40 000 words)
 - a. foreign students, FL
 - b. heritage French-Russian speaking students, HL
2. **French corpus** with French as L1 language
(16 500 words)
 - c. native French-speaking students, FrL1

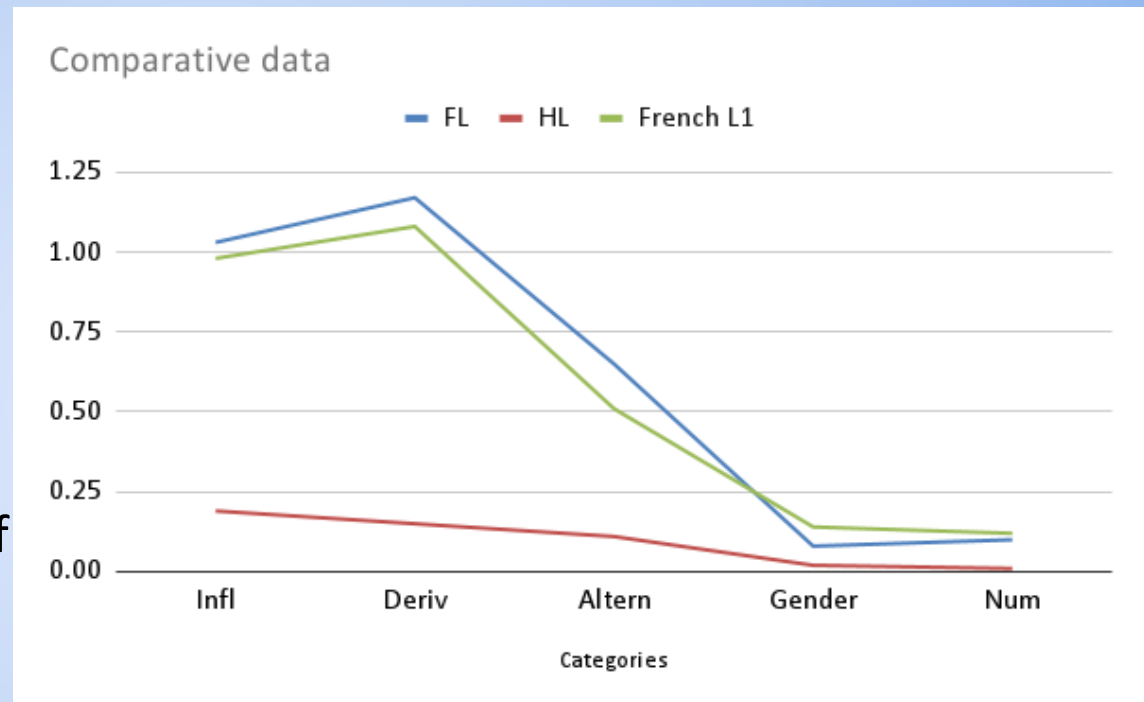
Step 2: Corpus Annotation

- language specific
- no previous deep studies on Russian learner corpus annotations
- We have selected **6 main categories** of students' errors:
 - **inflections, alternation, derivation, missing morphemes**
 - *show gaps in knowledge of Russian morphological system*
 - and **gender and number**
 - *reveal lack of acquisition of morphological features of nouns in gender (masculine, feminine, and neutral) and in number (singularia and pluralia tantum)*

Step 3: Quantitative Processing

- **Specific Observation:**

The most problematic morphological issue for French students is nominal and verbal **inflections**, as well as grammatical forms with **alternations** (nouns, verbs, prepositions). A large number of errors also involves **derivation**.

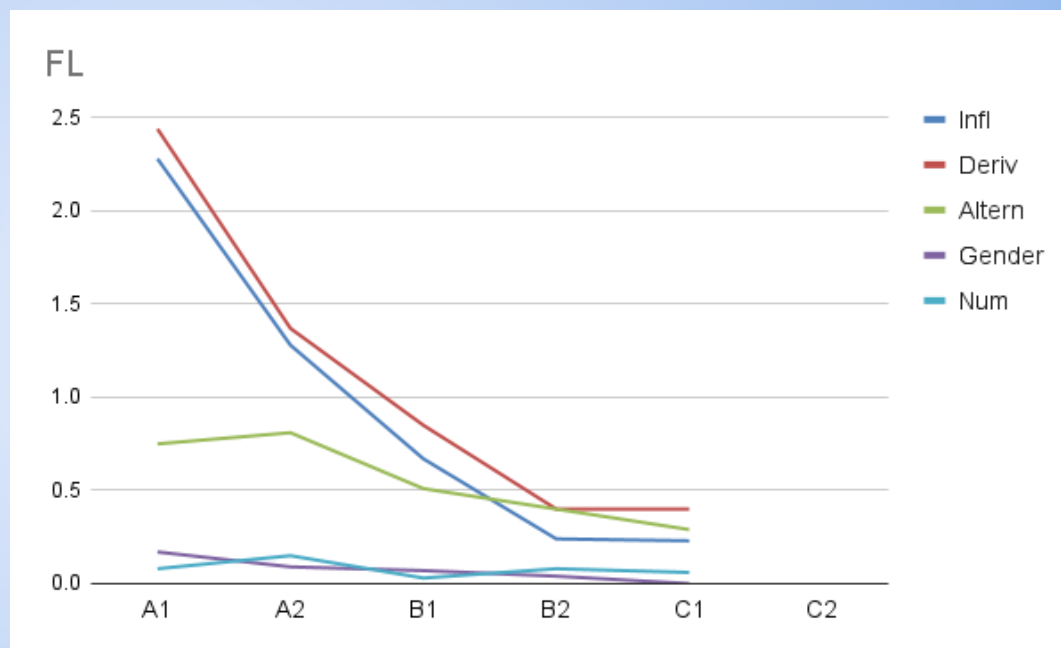


- **General Observation:**

Comparative data shows also that the L1 French corpus and the L1 multi-lingual French corpus (FL) present the same patterns. So, the native language has the same value in terms of linguistic transfer as a language of linguistic surrounding (French).

Step 4: Qualitative Analysis

- in progress
- involves **description** of specific problems with a high rate of errors in Russian L2 writing level by level:
 - noun / adjective and verbal inflections
 - gender agreement (interlinguistic factor, French transfer)
 - ...
- identifies **causes** of errors (intralinguistic, interlinguistic, and/or extralinguistic)
- and entails **comparative analysis with L1** acquisition of morphology



Conclusion and Further Research

This study demonstrates the value of quantitative analysis of a manually annotated corpus, which reveals general trends in the acquisition of Russian and describes learners' difficulties level by level within specific linguistic categories.

Analysis of the corpus data leads us in two possible directions:

- Instruction:
 - development of effective **teaching methods** focused on specific morphological issues
- Research:
 - understanding of **acquisitional processes** of Russian L2 morphology in a French environment
 - contribution to the **Second Language Writing** domain and to the **Learner Corpus Research**