# Frequent misspelling affects visual word processing: evidence from Russian

D.Chernova*, S.Alexeeva*, E.Kovalenko**, N.Slioussar*,**

*St.Petersburg State University

**Higher School of Economics

d.chernova@spbu.ru

**Introduction.** Most studies of spelling errors look at language production. Errors are seen as *a result* of a word's weak orthographic representation in an individual mind (see e.g. the Lexical Quality Hypothesis (Perfetti 2007)). However, Rahmanian and Kuperman (2019) suggested that frequent misspellings may also have an effect on processing — even if one knows how to spell a word, repeated exposure to incorrect spellings blurs its orthographical representation and weakens the connection between form and meaning. They confirmed this hypothesis in two experiments on English, an eye-tracking and a lexical decision study: in both studies, target words' reading times were affected by the incidence of spelling errors. It may be interesting to explore this issue in a cross-linguistic perspective. In this study, we turn to Russian, which differs from English by a more transparent orthography. In Experiment 1 we examine the robustness of orthographic representation and in Experiment 2 we check its influence on lexical access.

**Experiment 1** (N=166)**.** We selected 44 words that are frequently misspelled according to the Russian National Corpus (www.ruscorpora.ru), which mostly includes edited texts, and the social media subcorpus of the Taiga Corpus (Shavrina, Kurmachova, n.d.), which includes unedited texts. We used the error detection task: these words were presented in isolation in one of two conditions (either with or without a spelling error — 50% correctly and 50% incorrectly spelled words were distributed across two experimental lists). For every word, participants were asked to determine whether it is spelled correctly or not by pressing one of the two keys (RTs were not recorded).

**Results and discussion.** We used GLMMs for statistics. Responses to correctly and incorrectly spelled words were analyzed separately (the task of finding a particular error is different from making sure that no errors are present). Only the former are considered here to draw parallels with the Experiment 2 and Rahmanian and Kuperman's (2019) study. The frequency of the word and the relative frequency of its misspelled occurrences significantly influenced the number of incorrect responses. The first result is expected, while the second confirms the idea suggested by Rahmanian and Kuperman (2019) from a different perspective: not only it takes longer to read frequently misspelled words, it is also more difficult to decide whether they are spelled correctly.

**Experiment 2** (N=40)**.** We selected 30 words from the materials of Experiment 1. For every selected word, we found a pair that is matched for length and frequency, but is rarely misspelled due to its orthographic transparency. We used a lexical decision task, presenting these 60 words in the correct spelling, as well as 60 nonwords constructed by changing one letter in existing Russian words. In every trial, a letter string was presented on the screen for 500 ms, followed by a blank screen for additional 500 ms (if the participant responded earlier, the next trial started).

**Results and discussion.** We used LMMs for statistics. Firstly, the word type factor was significant: it takes more time to recognize a frequently misspelled word, which replicates the results by Rahmanian and Kuperman (2019) in a new language. Secondly, the interaction between the word type factor and the frequency factor was significant: the effect of misspellings was more pronounced for the words of lower frequency, contrary to the hypothesis put forward by Rahmanian and Kuperman (2019) (their own results for the frequency factor were inconclusive). We can conclude that high frequency words have more robust representations that resist blurring due to spelling errors more efficiently than low frequency ones. Finally, we conducted a separate analysis showing that the number of incorrect responses in Experiment 1 correlates with RTs in Experiment 2. Thus, whether we consciously try to find an error or simply read words (Rahmanian and Kuperman's eye-tracking experiment was especially telling in this respect because target words were presented in carrier sentences), orthographic representations blurred due to exposure to frequent misspellings make the task more difficult.

**Acknowledgements.** The study was supported by RFBR № 18-00-0646 (18-00-00640)

**References**

Perfetti C.A. (2007). Reading ability: Lexical quality to comprehension. *Scientific studies of reading, 11*, 357–383.

Rahmanian S., Kuperman V. (2019). Spelling errors impede recognition of correctly spelled word forms. *Scientific Studies of Reading, 23*, 24–36.

Shavrina T., Kurmachova Y. (n.d.), Taiga Corpus. An open-source corpus for machine learning, available at: https://tatianashavrina.github.io/taiga_site/.