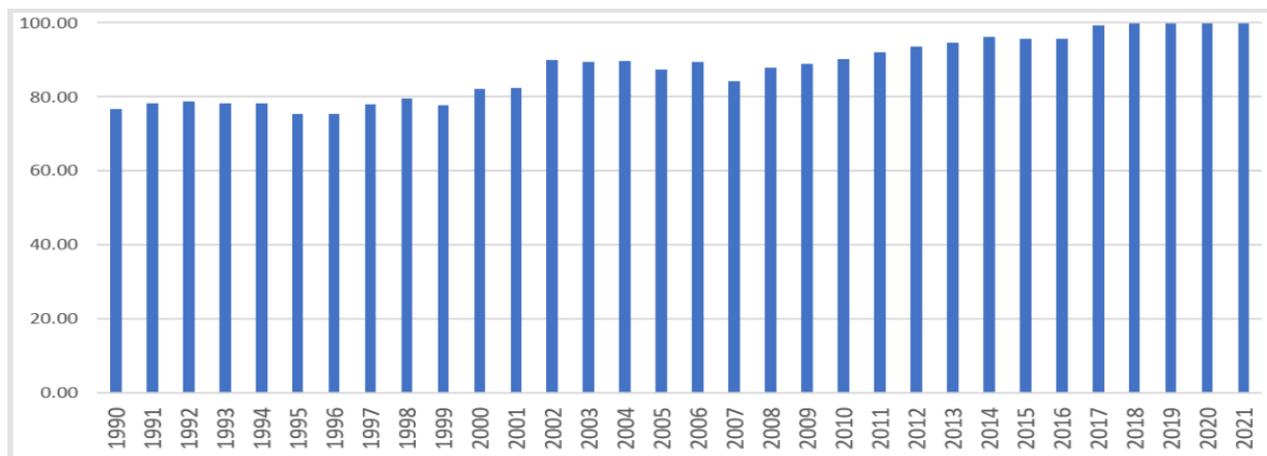


A Ukrainian-Russian Code-Switching Corpus of Ukrainian Parliamentary Sessions Transcripts

Maria Shvedova (Lviv Polytechnic, University of Jena), Olha Kanishcheva (University of Jena)
WG1

Bilingual Ukrainian-Russian parliamentary discourse → *Monolingual Ukrainian*

The Ukrainian language prevails in the corpus, and its share was increasing over the years: from a minimum of 76% in 1995 to 100% in 2018-2020.



- 70 M tokens
- 1990-2021
- Manual annotation by party, region, age
- Automatic language annotation (Lingua-py)
- Study of the correlation between language and political orientation
- Study of code-switching types, e.g., triggered code-switching (red for Russian, blue for Ukrainian):
Dear Vladimir Mikhailovich, Gennady Borisovich!
I would like to ask you to include in the list of objects the city of Kremenchuk and the city of Zolotonosha. These two cities are not on the list, and the problem is very acute in these two cities.
Thank you (Vasyl Havryluk, 2003).