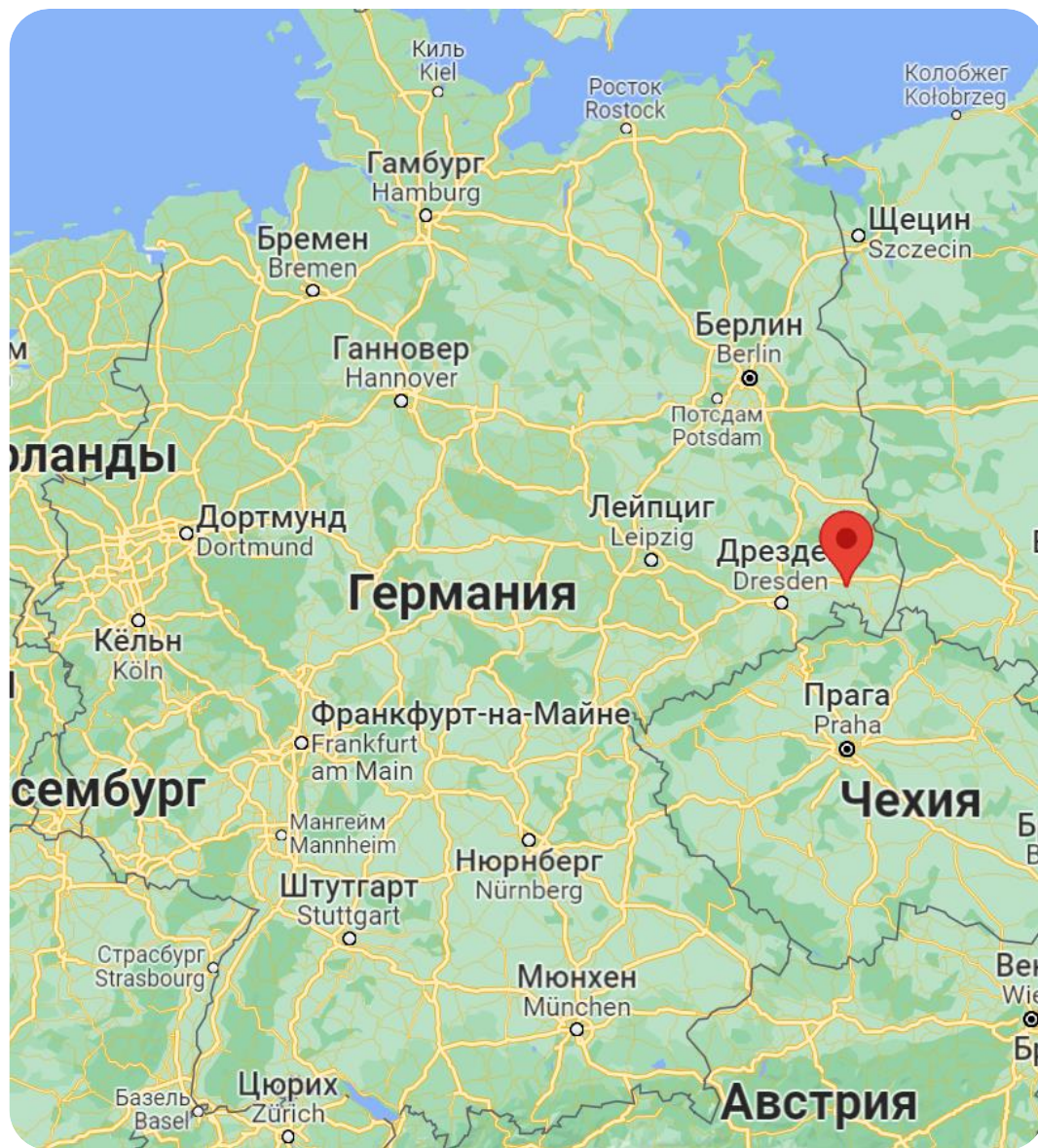


Верхнелужицкий

Игорь и Тарас



О языке

- **Верхнелужицкий язык** — один из двух языков лужицких сербов, распространённый в исторической области Верхняя Лужица на востоке Германии. Относится к лужицкой подгруппе западнославянских языков.
- Число говорящих — около 20 000 человек.

Почему Верхнелужицкий?

- Хорошо описанная и понятная грамматика
- Отсутствие данных на UniMorph
- Личный интерес

Какой корпус мы брали

- http://www.statmt.org/wmt20/unsup_and_very_low_res/ (составлен Серболужицким институтом в Баутцене; параллельный с немецким языком)
- Из них скриптом извлекли уникальные токены для разметки

Какая использовалась грамматика

- <https://baltoslav.eu/hsb/> + The Slavonic Languages, Greville Corbett, Bernard Comrie
- Хорошо описано словоизменение для всех частей речей

Разметили 400 токенов

ИТОГИ

durje	durjemi	N;INS;PL
minimalny	minimalna	ADJ;NOM;FEM;SG
suma	suma	N;NOM;SG
projekt	projekta	N;GEN;SG
dobry	dobrze	ADJ;ACC;NEUT;SG
zhromadny	zhromadne	ADJ;ACC;NEUT;SG
džěło	džěło	N;ACC;SG
knježerstwo	knježerstwje	N;LOC;SG
njebyć	njeje	V;PRS;3;SG
spomóžny	spomóžne	ADJ;ACC;NEUT;SG
swój	swoju	ADJ;ACC;FEM;SG
wulkosć	wulkosć	N;ACC;SG
skedźbnjeć	skedźbnjeć	V;NFIN
podšmórnyć	podšmórny	V;PST;3;SG
chcycь	chcemy	V;PRS;1;PL
wy	wami	ADJ;INS;PL
pohłubšować	pohłubšić	V;NFIN
wón	wona	ADJ;NOM;SG
wudžěleć	wudžěla	V;PRS;3;SG
slědźerski	slědźerske	ADJ;ACC;NEUT;SG
nadawk	nadawki	N;ACC;PL
organizować	organizuje	V;PRS;3;SG
dalši	dalše	ADJ;NOM;SG

Интересные случаи?

- Пока читали предложения, поняли, что язык SOV