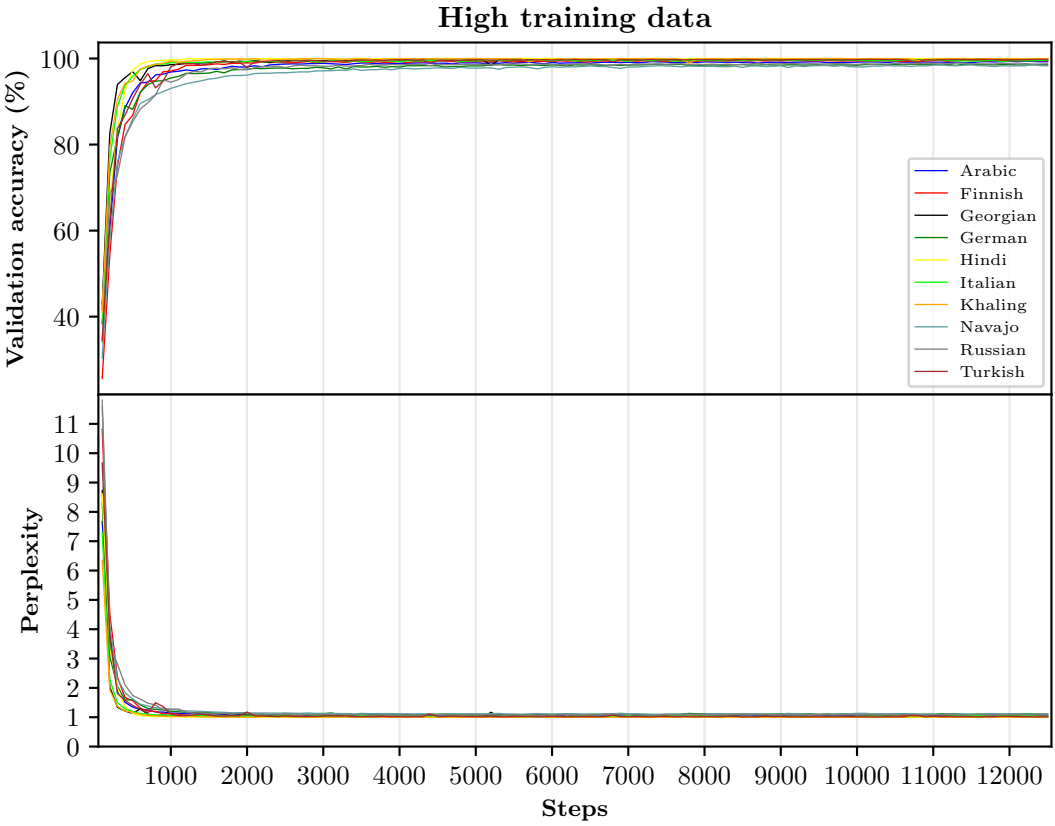
High training data
Validation accuracy (%)
100
80
60
40
Arabic
Finnish
Georgian
German
Hindi
Italian
Khaling
Navajo
Russian
Turkish
Perplexity
11
10
9
8
7
6
5
4
3
2
1
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
11000
12000
Steps