

new sensible word count

8000

6000

4000

2000

0

0.0

0.2

0.4

0.6

0.8

number of characters in the reference data

1e10

ar
bg
ca
cs
da
de
el
en
es
et
fa
fi
fr
he
hi
hr
hu
id
it
ko
lt
lv
ms
nl
no
pl
pt
ro
ru
sk
sl
sv
th
tr
uk
vi
zh

