5-gram Overlap versus Token Length for wcwidth **Match Type** 0 0.9 0.8 0.7 -**2-gram Overlap** 0.3 0.2 0 0.1 -

0

Length (tokens)

250

300

350

400

450

200

0.0

50

100

150

Exact Match No Match