

Corpus	Lines	Tokens
Training	10,313	218,021
Development	1,368	29,484
Test	1,371	30,263