PoS Within-Language Accuracy Drop 96.3% Korean 64.3% 32.0 -anguage Dataset 94.1% Turkish dev 74.9% **19.2** test 97.2% Arabic 76.7% **20.5** 30 10 20 40 50 60 70 80 90 100

Accuracy (%)