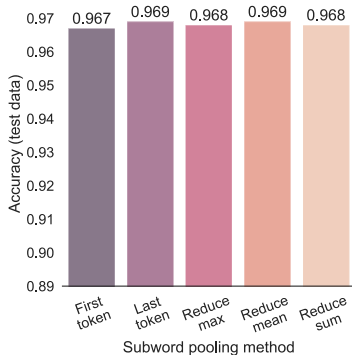
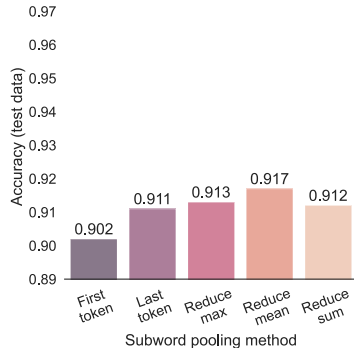


(a) Overall accuracy



(b) Accuracy on single-subword tokens



(c) Accuracy on multi-subword tokens

Figure 1: Effect of subpooling method on the accuracy of the POS tagger on the testing dataset. Note that accuracy should start from 0 but this scale is chosen to highlight the subtle differences between the taggers.