Propotion of different tokenization results (mBERT)

