Stand. deviation per language measured in accuracy (k=1) in different setups.