Setup info:

wait-k: 2, output rate: 2

pred. prob. threshold: 0.8

max prev. context size: 5

future-context token size: 1