Mean k Chosen at each Position during Generation with Nucleus Sampling b0.96-nowordcond p0.96-1wordcond Position in sequence