Train Loss (token vs no token) token 3.5 no token 3.0 2.0 1.5 50 25 75 100 125 150 175 Step