

Tokens		baseline	reinforced	150 steps	250 steps	350 steps	450 steps	550 steps	650 steps	750 steps	850 steps	950 steps	1050 steps	1150 steps
	Luke	0.5525	0.6213	0.1722	0.0750	0.0666	0.0614	0.0629	0.0669	0.0705	0.0709	0.0716	0.0719	0.0721
	the	0.2221	0.2598	0.1211	0.1359	0.1454	0.1525	0.1465	0.1536	0.1432	0.1367	0.1313	0.1336	0.1338
	up	0.0379	0.0042	0.0177	0.0231	0.0248	0.0256	0.0217	0.0218	0.0191	0.0186	0.0181	0.0183	0.0182
	him	0.0348	0.0011	0.1162	0.1101	0.1384	0.1622	0.1859	0.2080	0.2169	0.2225	0.2275	0.2334	0.2348
	his	0.0310	0.0125	0.0545	0.0510	0.0608	0.0691	0.0715	0.0758	0.0754	0.0735	0.0755	0.0768	0.0765
	for	0.0159	0.0097	0.0433	0.0487	0.0512	0.0560	0.0565	0.0617	0.0577	0.0579	0.0568	0.0570	0.0571
	hold	0.0155	0.0288	0.0196	0.0334	0.0340	0.0346	0.0341	0.0324	0.0313	0.0313	0.0310	0.0309	0.0306
	its	0.0095	0.0119	0.0234	0.0244	0.0210	0.0185	0.0185	0.0176	0.0160	0.0155	0.0160	0.0161	0.0161
	a	0.0095	0.0042	0.0142	0.0290	0.0308	0.0293	0.0264	0.0269	0.0238	0.0221	0.0208	0.0209	0.0210
	at	0.0062	0.0024	0.0090	0.0199	0.0339	0.0442	0.0492	0.0538	0.0567	0.0583	0.0575	0.0584	0.0587