

Model	Valid. Perp.	Test Perp.	Sec./ Batch
Baseline	54.1	37.7	.45
Multi-Token ( $L = 1$ )	54.2	37.4	.82
Multi-Token ( $L = 2$ )	53.9	36.4	4.85
Multi-Emb ( $E = 2$ )	<b>53.8</b>	<b>35.2</b>	2.53