

SYSTEMS	Voc.	En-Fr
Existing RNN systems		
LSTM (6 layers)	80K	31.50
LSTM (6 layers + PosUNK)	80K	33.10
Deep-Att	80K	37.70
Deep-Att + PosUnk	80K	<b>39.20</b>
GNMT WPM-32K	80K	38.95
DeepLAU + PosUNK trained on 12M data	80K	35.10
GNMT WPM-32K, HyperLSTM	80K	<b>40.03</b>
Existing Other systems		
ConvS2S (15 layers) + BPE-40K	40K	40.46
Transformer(base)	-	38.10
Transformer(big)	-	<b>41.00</b>
Our system		
RNN + BPE-40K	40K	38.19
MCE + BPE-40K	40K	38.80