Sutskever+ (2014) × Bahdanau+ (2014)

	Sutskever et al.	Bahdanau et al.
vocabulary	160k enc, 80k dec	30k both
encoder	4× LSTM, 1,000 units	bidi GRU, 2,000
decoder	4× LSTM, 1,000 units	GRU, 1,000 units
word embeddings	1,000 dimensions	620 dimensions
training time	7.5 epochs	5 epochs