Hyperparameter Name	Description	Values Tested
rand_reset	whether activations in the model reset to a random state (True), or to zero (False) after each word	**True**, False
lr	initial learning rate	0.1, 0.5, **1.0**, 2.0
anneal_factor	amount by which to anneal learning rate, if no improvement found	0, **0.25**, 0.5, 1.0
patience	number of training epochs to wait for validation loss to improve before updating weights	**0**, 2, 4
dropout	proportion of weights to keep fixed	0, **0.2**, 0.5
epochs	number of epochs (complete passes through the data) to train for	**25**, 50, 100
nlayers	number of recurrent layers	1, **2**, 4
nhid	number of nodes in each recurrent layer	128, 256, **512**, 1250

Table A1: Particulars of hyperparameter testing. Hyperparameters were optimized for speed and likelihood assigned to the validation set. Optimal parameters for the validation set are bolded, and were used in the experiments reported here.