

Global vs local attention

System	Perplexity	BLEU
global (location)	6.4	19.3
global (dot)	6.1	20.5
global (mult)	6.1	19.5
local-m (dot)	>7.0	x
local-m (mult)	6.2	20.4
local-p (dot)	6.6	19.6
local-p (mult)	5.9	20.9

Luong et. al. Effective Approaches to Attention-based Neural Machine Translation, 2015.