	Configuration	BLEU (Δ)	BP	$\mathcal{L}_1^{\text{GPU}}$ (Speed-up)		$\mathcal{L}_1^{\text{CPU}}$ (Speed-up)	
AT	*big* (teacher)	21.07	0.920	345 ms	1.0 $\times$	923 ms	1.0 $\times$
	base	18.91	0.908	342 ms	1.0 $\times$	653 ms	1.4 $\times$
	base (12-1)	15.47	0.806	152 ms	2.3 $\times$	226 ms	4.0 $\times$
	base (12-1) + KD	18.76	0.887	145 ms	2.4 $\times$	254 ms	3.6 $\times$
NAT	KD + CTC	16.93 (+0.00)	0.828	17.3 ms	19.9 $\times$	84 ms	11.0 $\times$
	KD + CTC + VAE	18.73 (+1.80)	0.862	16.4 ms	21.0 $\times$	83 ms	11.1 $\times$
	w. *BeamSearch20*	19.80 (+2.87)	0.958	28.5 ms	12.1 $\times$	99 ms	9.3 $\times$
	w. *BeamSearch20 + 4-gram LM*	**21.41 (+4.48)**	0.954	31.5 ms	11.0 $\times$	106 ms	8.7 $\times$
	w. *NPD5*	18.88 (+1.95)	0.866	34.9 ms	9.9 $\times$	313 ms	2.9 $\times$
	w. *NPD5 + BeamSearch20 + 4-gram LM*	**21.84 (+4.91)**	0.962	57.6 ms	6.0 $\times$	284 ms	3.2 $\times$