	WMT14		IWSLT14		
Models	En-De	De-En	De-En	Latency	Speedup
Non-autoregressive models					
FT [6]	17.69 (5.76)	21.47 (5.55)	/	$39ms^{\dagger}$	$15.6\times^{\dagger}$
FT [6] (rescoring 10)	18.66 (4.79)	22.41 (4.61)	/	$79ms^{\dagger}$	$7.68\times^{\dagger}$
FT [6] (rescoring 100)	19.17 (4.28)	23.20 (3.82)	/	$257ms^{\dagger}$	$2.36\times^{\dagger}$
Non-autoregressive models (Ours)					
NART	20.27 (7.14)	22.02 (9.27)	23.04 (10.22)	$26ms^{\ddagger}$	$14.9\times^{\ddagger}$
NART (rescoring 9)	24.22 (3.19)	26.21 (5.08)	26.79 (6.47)	$50ms^{\ddagger}$	$7.74\times^{\ddagger}$
NART (rescoring 19)	24.99 (2.42)	26.60 (4.69)	27.36 (5.90)	$74ms^{\ddagger}$	$5.22\times^{\ddagger}$