system	$EN\rightarrow DE$		$DE\rightarrow EN$	
	dev	test	dev	test
baseline	22.4	26.8	26.4	28.5
+synthetic	25.8	31.6	29.9	36.2
+ensemble	27.5	33.1	31.5	37.5
+r2l reranking	28.1	34.2	32.1	38.6