Two-Step Attempts (WMT10) 2/2

| Training Sents | | Vanilla | | Two-Step | | Diff |
Parallel	Mono	BLEU	SemPOS	BLEU	SemPOS	B.S.
126k	126k	10.28 ± 0.40	29.92	10.38 ± 0.38	30.01	↗↗
126k	13M	12.50 ± 0.44	31.01	12.29 ± 0.47	31.40	↘↗
7.5M	13M	14.17 ± 0.51	33.07	14.06 ± 0.49	32.57	↘↘

Manual micro-evaluation of ↘↗, i.e. 12.50 ± 0.44 vs. 12.29 ± 0.47:

	Two-Step	Both Fine	Both Wrong	Vanilla	Total
Two-Step	**23**	4	8	-	**35**
Both Fine	7	14	17	5	43
Both Wrong	8	1	28	2	39
Vanilla	-	3	7	**23**	33
Total	**38**	22	60	30	150

- Each annotator weakly prefers Two-step
  - but they don't agree on individual sentences.