Zero-shot MT

Firat et al. 2016 (two papers)

- *Finetuning*: what if we use a small amount of parallel data in this setting?

Pivot	Many-to-1		Pseudo Parallel Corpus				True Parallel Corpus			
			1k	10k	100k	1m	1k	10k	100k	1m
Single-Pair Models		Dev	–	–	–	–	–	–	11.25	21.32
		Test	–	–	–	–	–	–	10.43	20.35
√	No Finetuning		Dev: 20.64, Test: 20.4				–			
		Dev	0.28	10.16	15.61	17.59	0.1	8.45	16.2	20.59
		Test	0.47	10.14	15.41	17.61	0.12	8.18	15.8	19.97