

Simple Data Mixing

Do we really need separate encoders and decoders?

... simply feed in various language pairs:

Source Sent 1 (De)	2en versetzen Sie sich mal in meine Lage !
Target Sent 1 (En)	put yourselves in my position .
<hr/>	
Source Sent 2 (En)	2nl I flew on Air Force Two for eight years .
Target Sent 2 (Nl)	ik heb acht jaar lang met de Air Force Two gevlogen .

- The model of the same size will learn both pairs.
- Hopefully benefiting from various similarities.
- Risk of catastrophic forgetting.

See Johnson et al. (2016) or Ha et al. (2017).