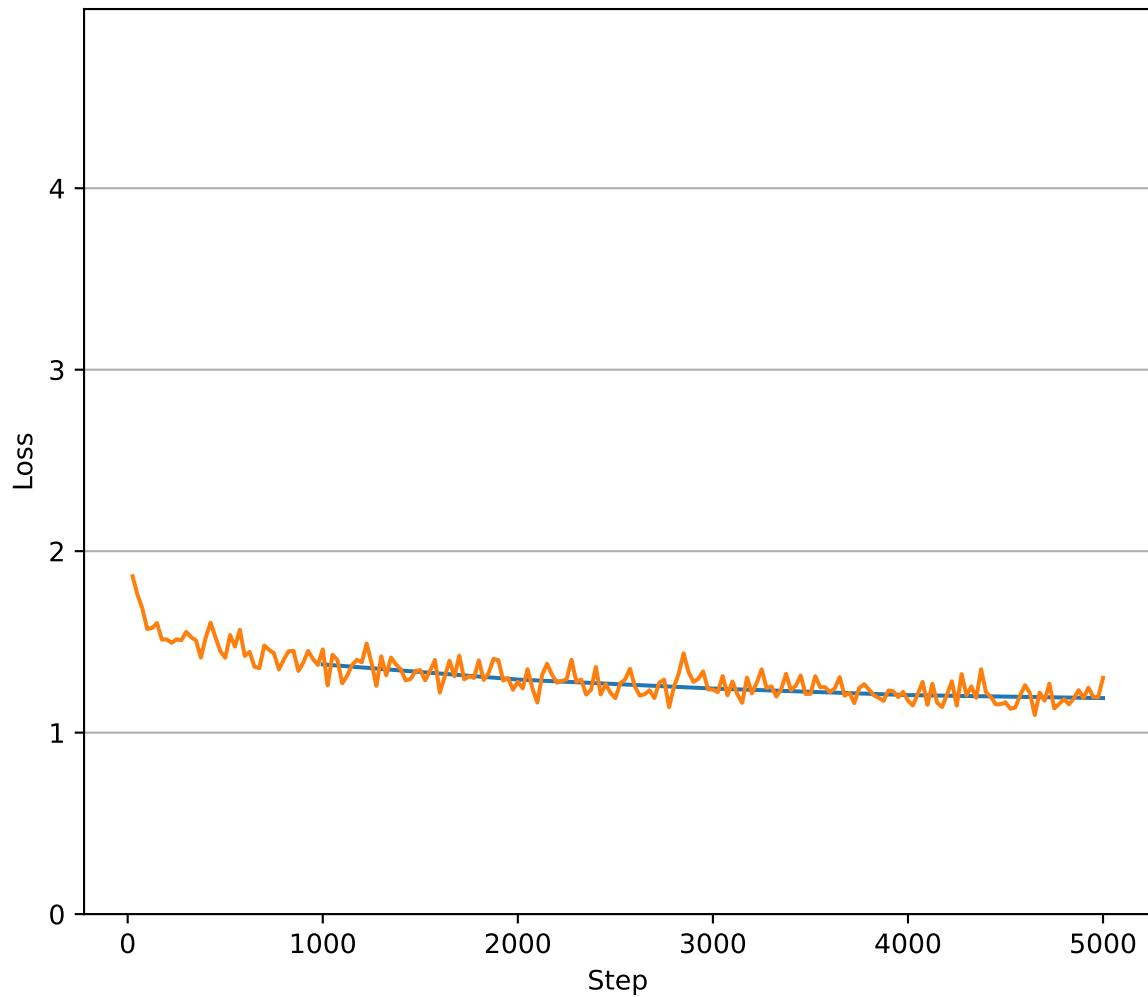


Loss over time for pooled dialect data

With pre-training



Without pre-training

