



Monolingual Transfer Learning via Bilingual Translators for Style-Sensitive Paraphrase Generation



Tomoyuki Kajiwara Biwa Miura Yuki Arase

We tackle the low-resource problem (< 100k sent. pairs) in style transfer

Src: **Informal** | L00000VVVVVVVEEE this song SOOO Much!!!!!!
 Ref: **Formal** | **very much enjoy** this song.
 [Rao+ 2018] | **loovvvvvveee** this song so **Much**.
 [Niu+ 2018] | **really enjoy VVVVVVVVEEE** this song.
 Ours | **love** this song **very much**.

Example of Informal → Formal style transfer in GYAFC corpus.

Rule-based data augmentation

Multitask learning with style-sensitive machine translation

Independent of any manual annotation or a special dataset

Pre-training with a raw corpus → Fine-tuning with a parallel corpus

① Roundtrip Translation

Generate a pseudo-parallel corpus using a large-scale raw corpus. We get the paraphrases in the table below.

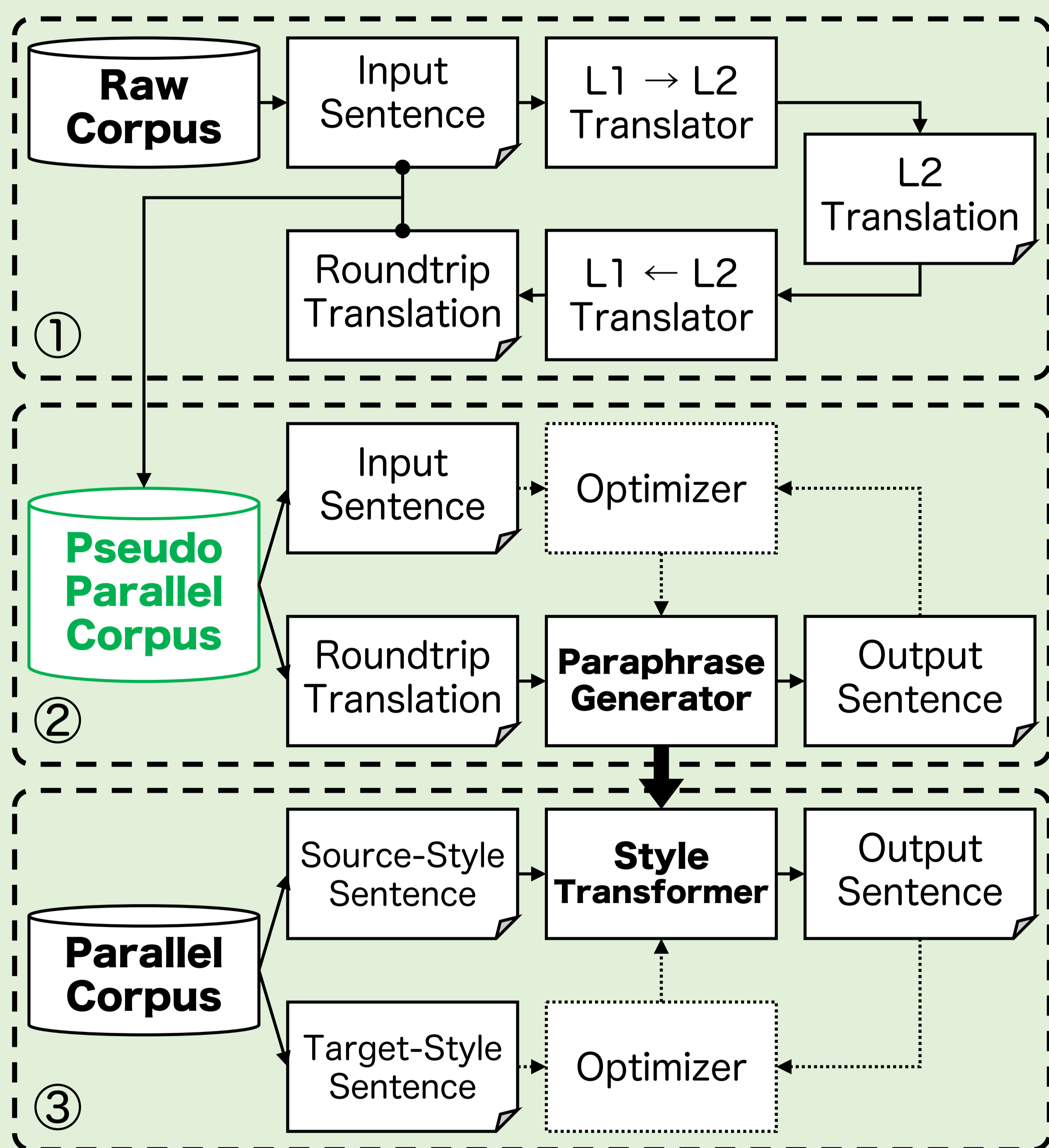
② Pre-training

Train a paraphrase generator that reproduces an input sentence from roundtrip translation. Note that this model has not yet learned about any style.

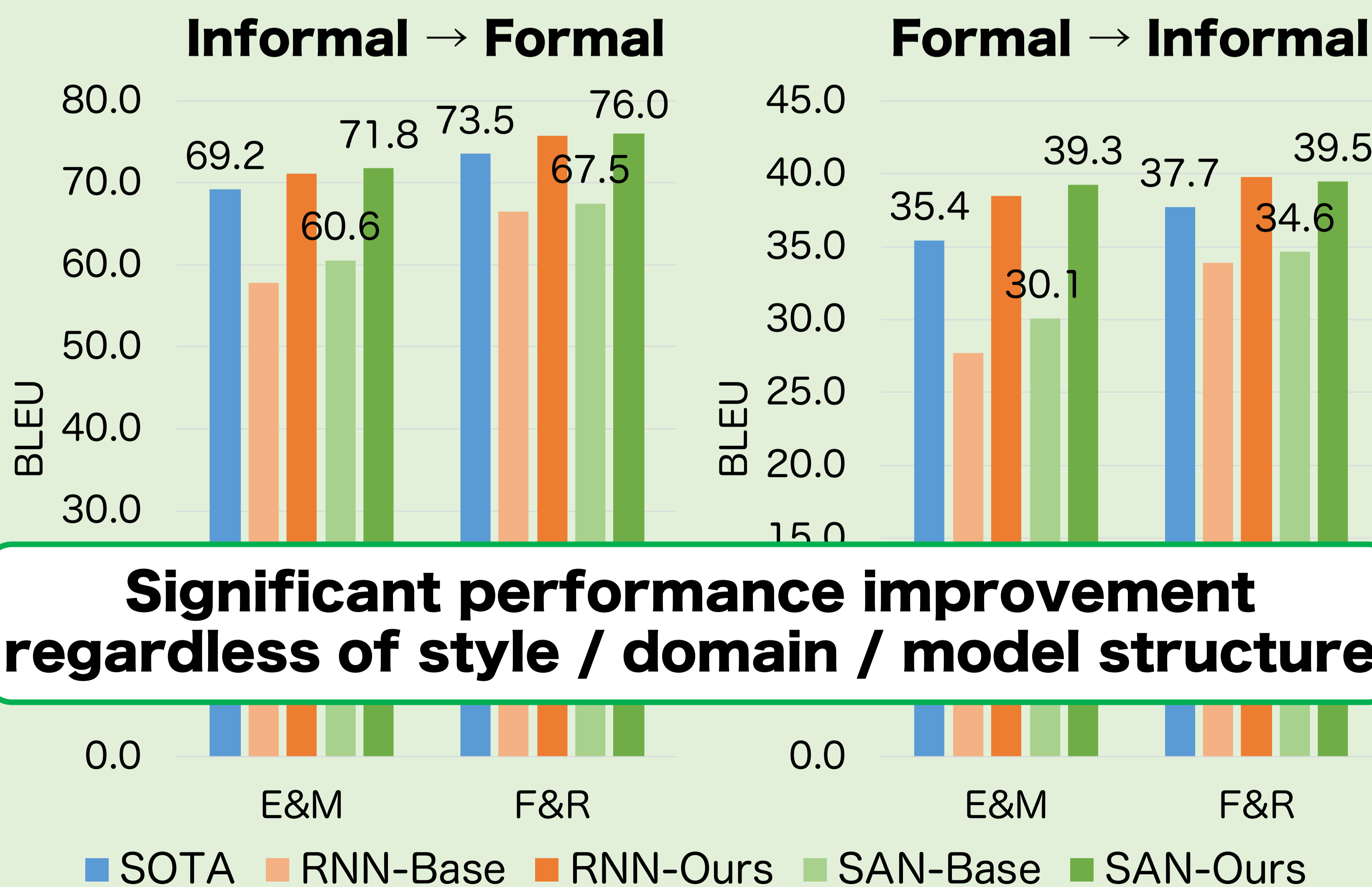
③ Fine-tuning

Train the model further using a parallel corpus of the target style. Pre-training allows the model to learn synonymy and grammaticality, so it can effectively learn style from such small-scale parallel corpus.

Input Sentence	Roundtrip Translation
I love watching the show.	I love to see the show.
Thanks for asking the question.	Thank you for the question.
The key to a successful relationship is good communication .	Good communication is the key to a successful relationship.

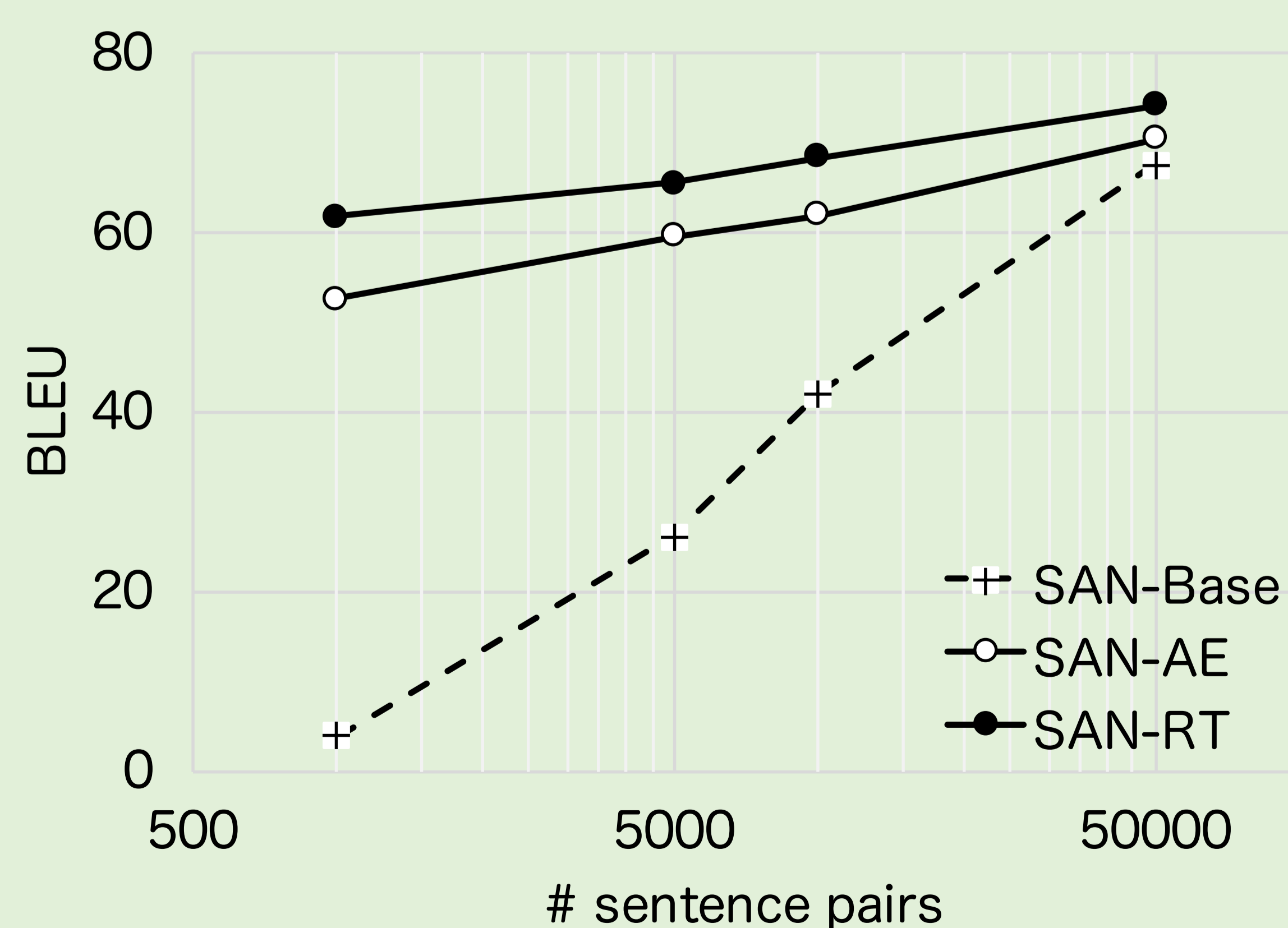


Successful style transfer with only a 1k parallel corpus



Significant performance improvement regardless of style / domain / model structure

Informal → Formal (F&R domain)



Pre-training based on **auto-encoding** is also effective if bilingual translators are not available.