

Chatbots can be made more polite using style transfer & synthetic data

Polite Chatbot: A Text Style Transfer Application

Sourabrata Mukherjee
Vojtěch Hudeček
Ondřej Dušek

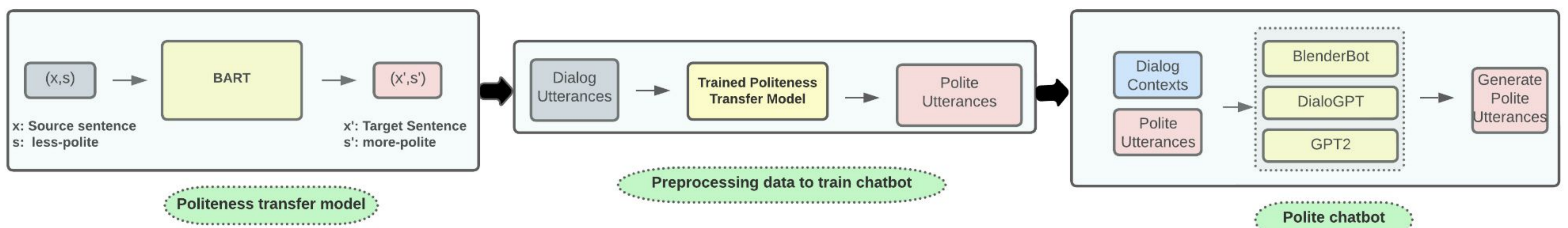
mukherjee@ufal.mff.cuni.cz
hudecek@ufal.mff.cuni.cz
odusek@ufal.mff.cuni.cz



CHARLES
UNIVERSITY



Our approach



1. Politeness transfer model:

- politeness markers: identified and removed from polite texts (TF-IDF-based)
- BART trained to reconstruct politeness markers

2. Synthetic polite chat data:

- generated by applying politeness transfer model

3. Polite dialogue model:

- BlenderBot, DialoGPT, and GPT-2 as the dialog models
- finetuned them using synthetic polite data for generating polite responses

Experiments

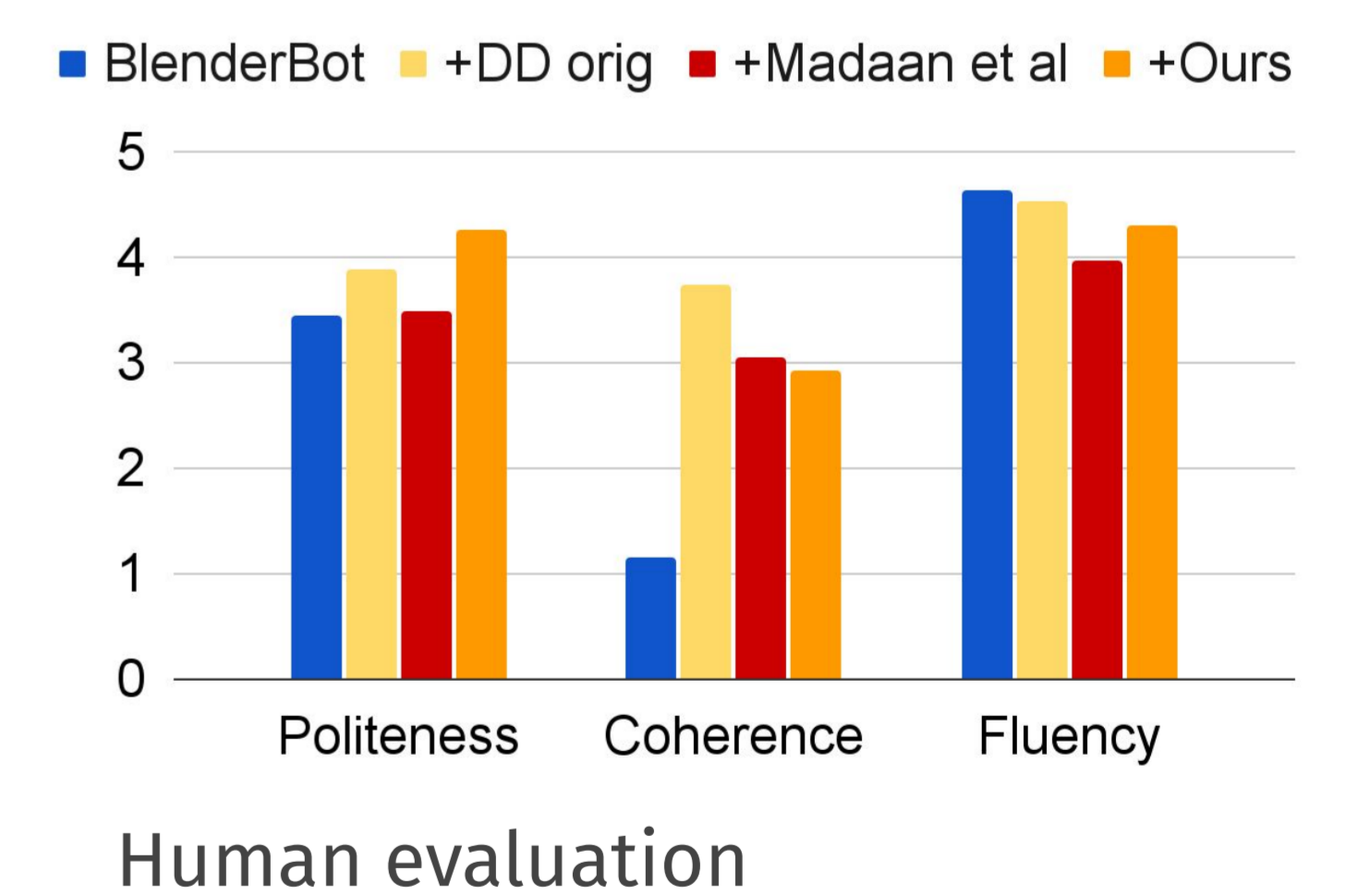
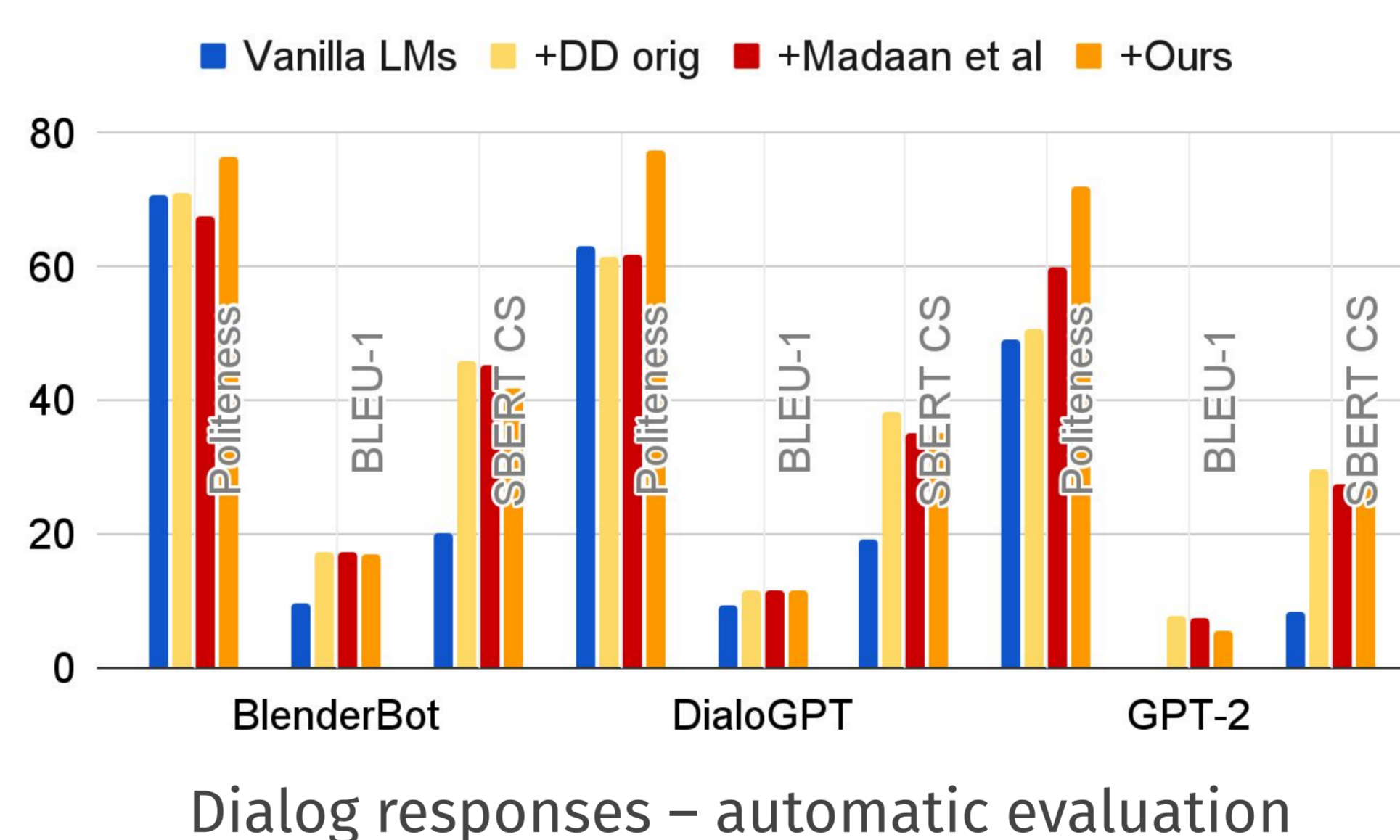
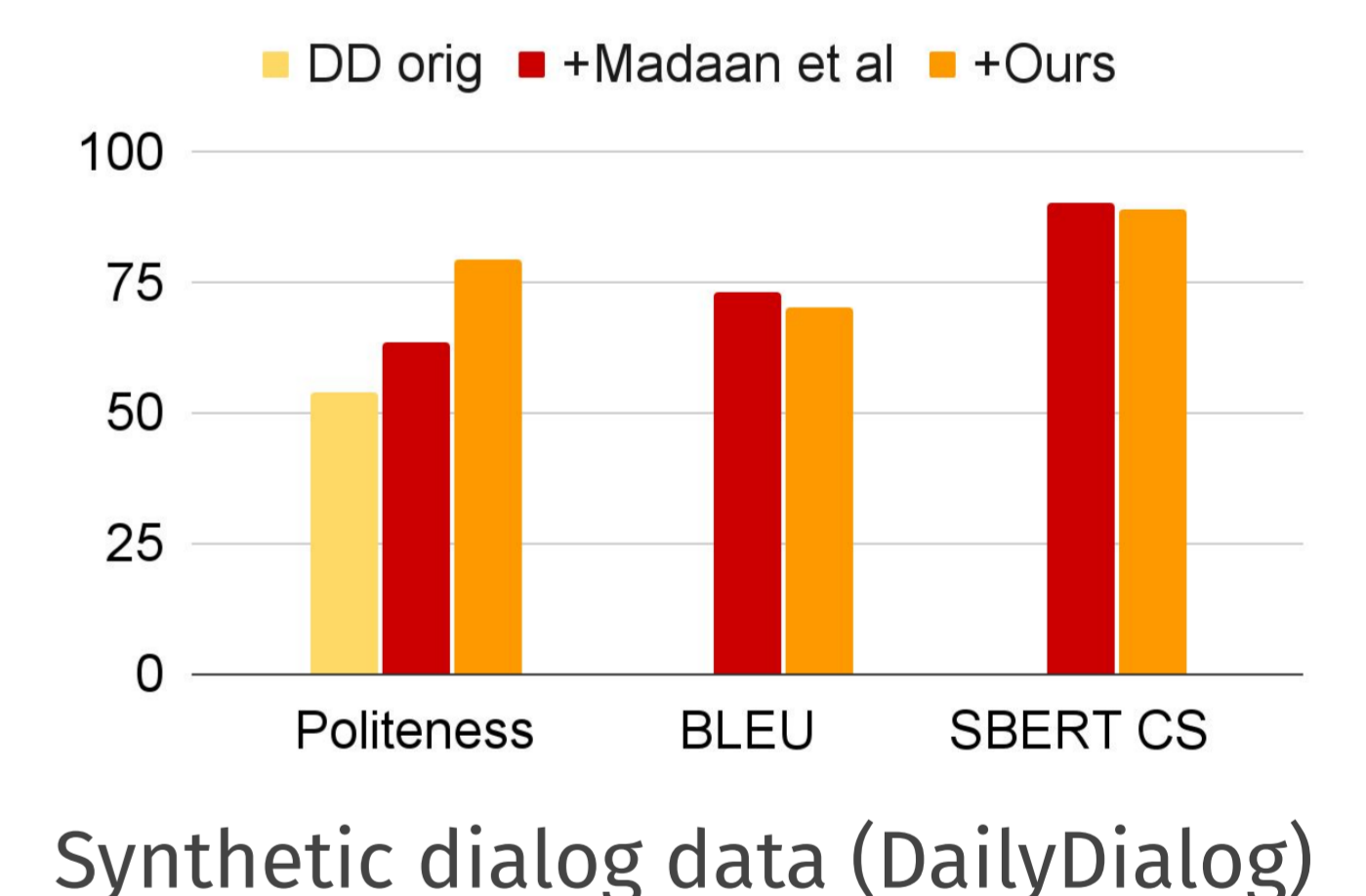
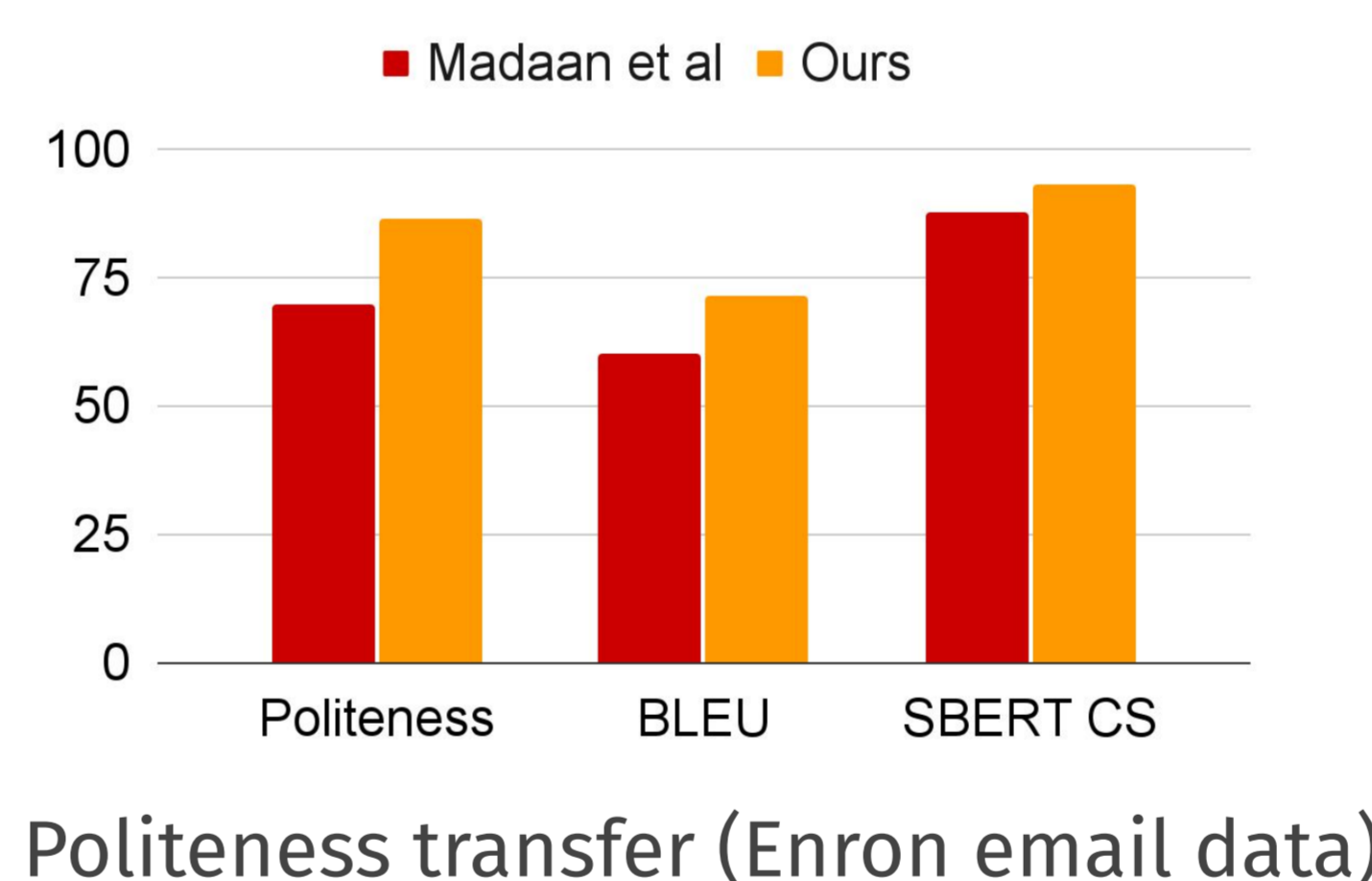
Data

- Politeness:
 - preprocessed & filtered Enron email sentences (Madaan et al., 2020)
- Dialog:
 - DailyDialog (DD): open-domain human-human dialogues (Li et al., 2017)

Metrics

- Automatic evaluation
 - politeness: politeness classifier
 - coherence: BLEU, SBERT cosine similarity (CS)
- Human evaluation
 - politeness, coherence, fluency
 - 1-5 Likert scales

Results



Conclusions

Resulting model is end-to-end, no post processing needed

Compared to baselines, we achieve:

- Improved politeness
- Comparable content similarity, fluency, coherence

Future work: other stylized response generation tasks

Sample outputs

Context : Nani Book Store. How can I help you? Do you have the 'The man and the Sea' by Ernest Hemingway?	
Ground Truth	Yes, we have one copy left.
Vanilla (no FT)	I would love to try it, but I don't think I'll be able to afford it.
DailyDialog (DD)	Yes, it 's on that shelf. I 'll get it for you right away.
DD + Madaan et al. (2020)	We do. Would you like a look at it?
DD + Ours	Yes, we do. Would you like to have a look at it? please let me know.

