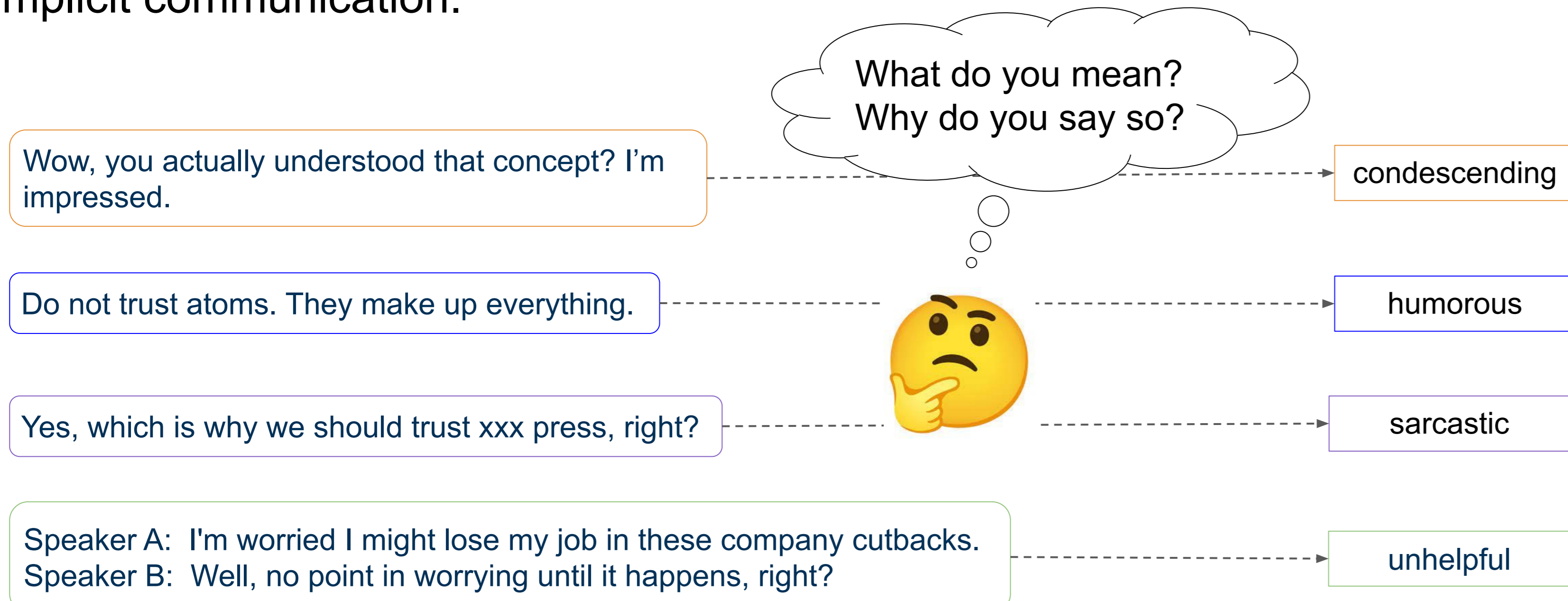


1. Introduction

Style Identification: The style of a text refers to unique ways authors select words and grammar to express their message. It can provide insights into social interactions and implicit communication.



Challenges:

- (1) The open-ended and ever-evolving nature of style motivates the need for zero-shot learning.
- (2) Zero-shot style classification remains a challenge for standard LLM prompting.
- (3) The ambiguity of style in language often poses a challenge for classification efforts.

2. Contributions

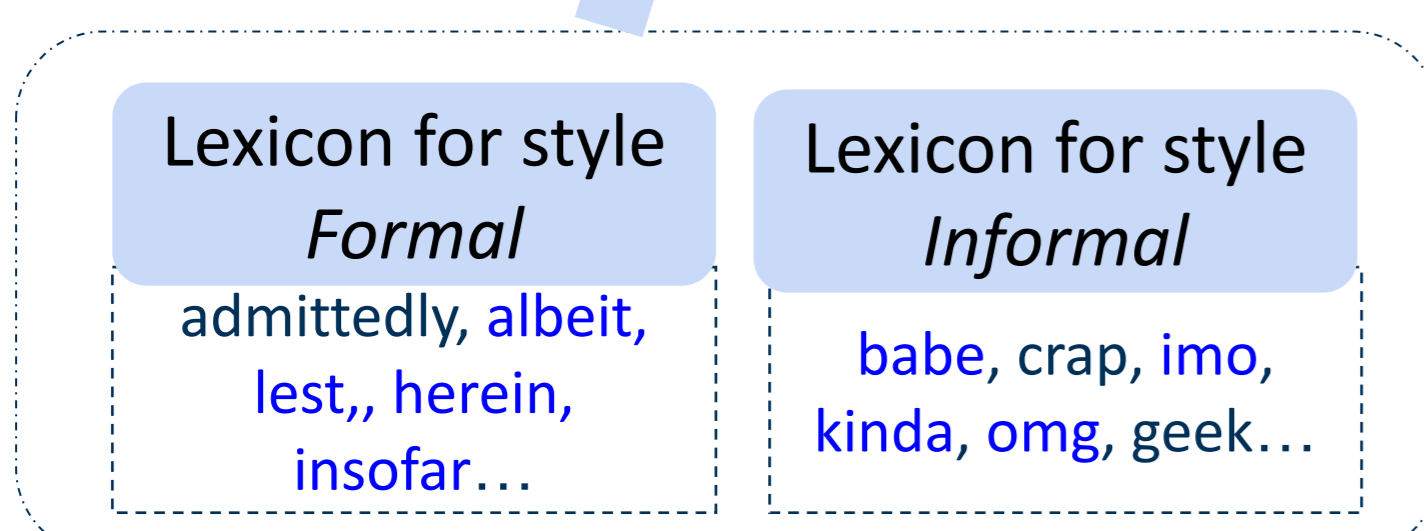
- Introduce **lexicon-based instructions**, a simple yet effective method for zero-shot style classification leveraging lexical knowledge in LLMs.
- Show **class randomization** can improve generalization capability of lexicon-instructed models significantly
- Provide a **benchmark**, featuring 13 established tasks and a synthetic dataset of 63 new tasks

3. Lexicon-Based Instruction

Method

Instruction: The task is to classify a sentence as "**formal**" if the style of the sentence is similar to the words "**albeit, lest, herein, insofar**" or as "**informal**" if the style... words "**babe, imo, kinda, omg**". Here is the sentence: "I think she is unvirtuous."

Add Lexicon



Key idea: Meta-tune the model on training styles using lexicon-based instructions.

5. Zero-Shot Learning

On 6 established datasets

Model	Meta-Tuned?	Instruction	Shakespeare	Romance	Humor	Country	Sarcasm	Age	Avg.
Flan-T5 _{base}	✗	Standard	33.36	33.33	33.33	43.15	33.33	33.92	35.07
	✗	+ Lex	49.95	51.30	48.66	35.34	49.40	49.02	47.28
Style-T5 _{base}	✓	Standard	33.31	43.57	36.43	19.86	33.37	35.75	33.72
	✓	+ Lex	55.10	78.98	60.56	49.09	49.25	50.80	57.30
Style-GPT-J	✓	Standard	58.16	87.82	33.33	53.11	44.10	35.25	51.96
	✓	+ Lex	56.76	83.99	55.86	44.97	48.84	47.47	56.32
LLaMA-2-Chat (7B)	✗	Standard	60.20	85.72	43.84	49.19	36.02	38.91	52.31
	✗	+ Lex	62.59	88.95	51.01	50.88	42.88	36.54	55.47
LLaMA-2-Chat (13B)	✗	Standard	61.99	97.00	47.42	17.96	43.26	48.16	52.63
	✗	+ Lex	63.49	95.00	55.15	24.41	44.66	53.88	56.10
LLaMA-2 (7B)	✗	Standard	42.13	64.41	37.38	48.27	48.84	37.13	46.36
	✗	+ Lex	50.21	77.86	45.44	49.86	47.72	47.63	53.12
Style-LLaMA (7B)	✓	Standard	40.91	41.65	48.88	48.92	49.02	49.80	46.53
	✓	+ Lex	59.03	88.97	57.64	51.52	50.83	50.53	59.75

On 20 synthetic tasks

	Standard	+ Lex
Random Baseline	36.65	
LLaMA-2-Chat (7B)	53.09	56.23
Style-LLaMA (7B)	46.25	58.71
Style-LLaMA+ (7B)	65.46	74.31
LLaMA-2-Chat (13B)	56.80	59.75

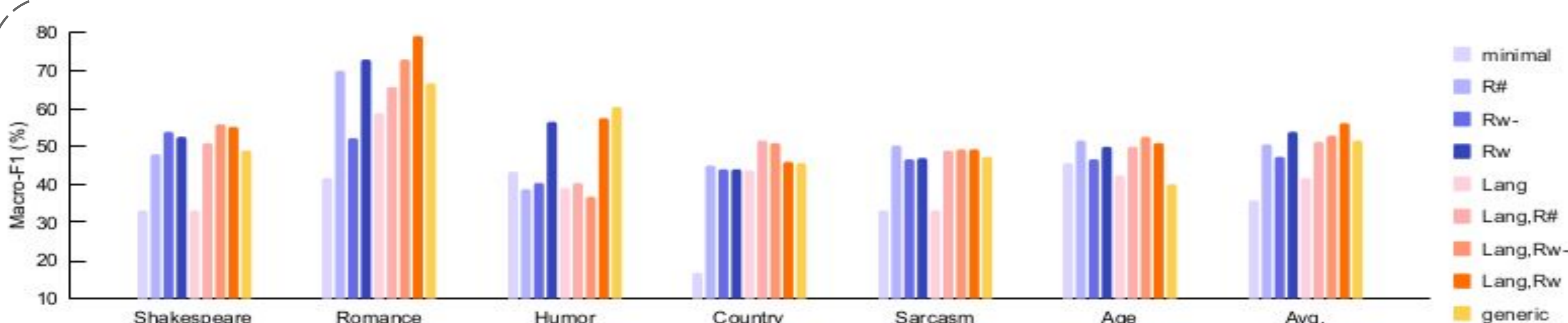
- **Lexicon-based instructions outperform standard instructions.**
- **Meta-tuning on style data with lexicons enhances the zero-shot performance compared to general instruction tuning.**

4. Class Randomization

Method

Instruction: The task is to classify a sentence as "**southnp**" if the style of the sentence is similar to the words "**albeit, lest, herein, insofar**" or as "**Atlanta**" if the style... words "**babe, imo, kinda, omg**". Here is the sentence: "I think she is unvirtuous."

- **Goal:** Improve the instruction-tuned models' ability to generalize to unseen styles.
- **Key idea:**
 - Replace meaningful style labels with **randomly selected identifiers**, in order to prevent models from memorizing training styles.
 - Encourage the model to generalize to interpret lexicons for new styles.



Class randomization matters in lexicon-based meta-tuning.