Pre-trained Language Models' Interpretation of Evaluativity Implicature: Evidence from Gradable Adjectives Usage in Context

Yan Cong yancong222@gmail.com https://yancong222.github.io/

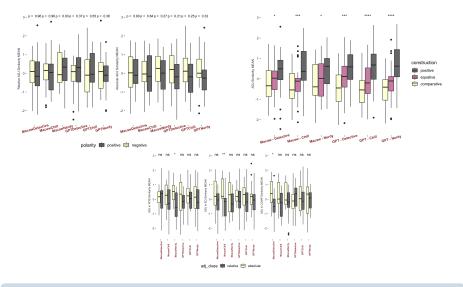
UnImplicit Workshop @ NAACL 2022

	Relative tall	Absolute full	
	POSITIVE tall NEGATIVE short	Positive full negative empty	
POS	less evaluative Maria is tall	more evaluative is full	
EQ	more evaluative Maria is as tall as	less evaluative is as full as	
COMP	exhibit the same EVAL is taller than vs. is more full than		

Table: 1 Distribution of EVAL implicature - human data: (1) POS is the most EVAL; (2) EVAL is sensitive to adjective class; (3) no EVAL difference in polarity (i.e., positive and its negative antonym) (Brasoveanu and Rett, 2018).

Prompts	GPT-3 response	Macaw response
The Detective reported to the Police Chief: "Maria is tall." What can the Chief conclude from this? [DETECTIVE]	The Police Chief can conclude that Maria is taller than average The Detective is reporting on the height of Maria.	. Maria is tall.
Arthur is clever, smart, quiet, awkward, nerdy, shy and geeky. What does he mean by saying "Maria is tall"? [Nerd]	Maria is taller than Arthur Maria is tall because she is taller than most people.	
Alex is chill, laid-back, relaxed, easy, cool, friendly, and outgoing. What does he imply by saying "Maria is tall"? [Chill]	Maria is tall implies that Maria is taller than most people. Maria is taller than Alex.	

Table: 2 Example LMs completions with POS as prompt. EQ and COMP were situated in the same prompt frame.



Take-home message: LMs results

- O Polarity: most of the LMs outputs align with the human data
- Constructions: POS was the most evaluative across different LMs and adjective types
- Adjective class: relative adjective class was less evaluative; almost always insensitive to constructions
- Socio-pragmatic persona: helped some LMs "understand" implicitness