

## Exploring the Agent-Relativity of Truth

Imagine a context where a speaker is justified in making a claim but its content does not correspond to the facts. Recent work in experimental philosophy and semantics has uncovered that English speakers tend to split on the truth value of sentences like "Joe might be in Boston" in such contexts (Knobe & Yalcin 2014, Phillips & Khoo 2019, Phillips & Mandelkern 2020). Subsequent empirical studies even suggest that truth-value judgments of simple declarative sentences like "Joe is in Boston" show surprisingly high variability (Reuter & Brun 2021, Ricciardi & Martin 2022). These findings pose a challenge to the commonly assumed view in formal semantics that what makes a sentence true or false is just its correspondence with the facts (**correspondence sense of "true"**). So, what underlies the variability in truth-value judgments? Reuter & Brun 2021 hypothesized that such variability is due to an inherent ambiguity of the term "true" between the correspondence sense and a **coherence sense**, according to which a sentence is true or false depending on whether its content coheres with the speaker's set of beliefs at the time when she utters the sentence. The study we present investigates the following question:

**Question:** What is the key determinant in activating a particular sense of "true" across contexts?

In this paper, we hypothesize that the key determinant is **agent-relativity**. More specifically, we predict that if the focus is on the sentence uttered by a speaker as in "Is it true that [sentence]?", people will tend to apply a correspondence sense of "true". In contrast, if the focus is on the agent making a statement as in "Is it true what A said?", people will be more inclined to apply a coherence sense of "true".

**Methodology.** We designed a two-response options questionnaire with a 2 x 5 design where participants first read one of two stories adapted from previous works (**see 1**) and then answered one of five questions (**see 2**). We recruited 400 participants from Prolific Academic who were randomly assigned in batches of 40 to one of the ten conditions.

### (1) The two stories read by participants

#### *Story A: Party*

Maria and Peter are students and meet up for a late dinner. Peter asks Maria whether Tom is at the party that they intend to go to after dinner. Maria answers that **Tom is at the party**. After all, Tom had told her that he would be at the party. When they arrive at the party, it turns out that Tom has changed his plans, and is not at the party.

#### *Story B: Boston*

Sally and George are meeting up in a cafe in the afternoon, talking about whether Joe is currently in Boston. Yesterday, Joe told Sally that he would have a job interview in Boston at 5 pm today and he would fly there early in the morning. So, Sally states: "**Joe is in Boston**". Just then, George gets an email from Joe. The email says that the job interview was canceled and that he is still in Berkeley. So George says: "No, he isn't in Boston. He is in Berkeley."

### (2) The five critical questions (A = speaker, S = sentence)

Question type	Label	Response Options
---------------	-------	------------------

Has A said the truth?	(A Said Truth)	Yes/No
Was A's answer true or false?	(A's Statement)	True/False
Is it true what A said?	(What A Said)	Yes/No
Is the underlined statement true?	(Pure Statement)	Yes/No
Is it true that S?	(Fact)	Yes/No

**Results.** We ran  $\chi^2$ -square tests to assess the impact of the two independent variables, *scenario* and *question type*, on participants' responses. For *scenario*, the analysis revealed a significant effect with  $\chi^2 = 13.33$ ,  $p < 0.01$ , Cramer's  $V = 0.183$ . Regarding *question type*, the test showed a significant effect with  $\chi^2 = 100.55$ ,  $p < 0.001$ , Cramer's  $V = 0.502$ . For Story A Party, the proportions of 'Yes/True' responses by question type were as follows: Truth (77.5%), A's statement (55%), What A Said (52.6%), Pure Statement (12.5%), and Fact (0%). For Story B Boston, the proportions of 'Yes/True' responses by question type were as follows: Truth (60%), A's Statement (32.5%), What A Said (25%), Pure Statement (17.5%), and Fact (2.5%). **See also Figure 1.**

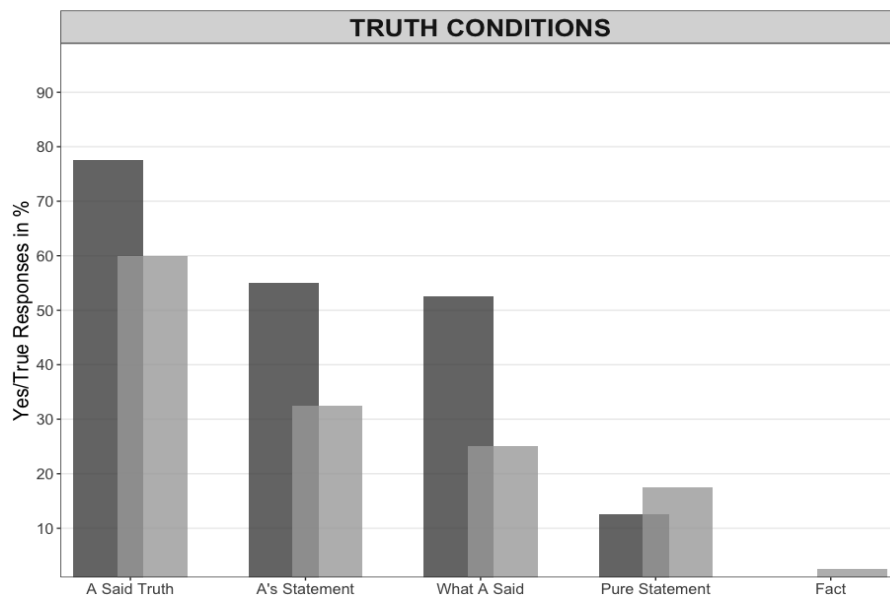


Figure 1. Yes/true responses in % of participants in the five question types tested. Dark grey columns show the results for the Party story, light grey columns for the Boston story.

**Conclusion.** In this work, we investigated what drives variance in truth-value judgments across different contexts. Our results show that when the task instructions are phrased in agent-relative terms like in "Has A said the truth" or "Was A's answer true or false?", a large proportion of people judge the sentence to be true; instead, when the task instructions focus on the sentence itself like "Is the underlined statement true?" or "Is it true that S?", most participants converge in judging the sentence to be false. We take these findings as suggesting that with task formulations focusing on the agent, many people seem to access the coherence sense of "true" whereas with task formulations focusing on the sentence people converge on the correspondence sense. Overall, our findings indicate the need for further investigations into how naive participants interpret the terminology employed in experimental semantic tasks. These investigations are crucial to test further hypotheses on the exact meaning of the two senses of "true", as well as possible pragmatic factors that drive people's responses.