

# *Believe* is Strong but Subjective: Experimental Evidence from Hedging

Todor Koev, Cory Bill, and Maryam Mohammadi

University of Konstanz

*Sinn und Bedeutung* 25, UCL and QMUL

Main Session

September 1-9, 2020

# Introduction

- The core structure of a *believe* clause:

(1)    I    believe    it is raining  
         agent    main predicate    prejacent

## Main question

What kind of certainty does *believe* ascribe to the agent regarding the prejacent?

- All scholars agree that *believe* conveys some sense of weakness.
- But what kind of weakness exactly?

# Two competing views

- **Hawthorne et al. 2016**

- *Believe* carries a weak/non-strong modal force.
- It is akin to the probability operator *likely*.

(2) I believe the Giants will win.

↪ “It is likely that the Giants will win.”

- **Koev 2019**

- *Believe* carries a strong modal force but a weak modal content.
- It conveys maximal but subjective certainty, thus implying lower objective certainty.

(3) I believe the Giants will win.

↪ “Personally, I am absolutely certain that the Giants will win (but I lack good evidence for it).”

## Hedging with *believe*

- How to distinguish between these two views?
- We will focus on the use of *believe* as a **hedge**.
  - (4) I believe the Giants will win, but I'm not sure they will.
- We will show experimentally that such data favors the latter strong-but-subjective view over the former weak view.
- Main evidence: The availability of hedging sentences is affected by certain grammatical and discourse factors.

# Preview

- The standard Hintikka semantics and strength
- Two views on the strength of *believe*
- Empirical evidence for each view
- Predictions about the hedging data
- Experiment 1 (grammatical factors)
- Experiment 2 (discourse factors)
- General discussion
- Subjectivity and epistemic modality

# The Hintikkan orthodoxy and the issue of strength

- *Believe* involves universal quantification over possibilities, stating that the prejacent is true in all of the agent's doxastic alternatives (Hintikka 1969).

$$(5) \quad \llbracket \text{believe} \rrbracket^i = \lambda p \lambda x . \forall i' \in \text{Dox}_{x,i} : p(i'),$$

where  $\text{Dox}_{x,i}$  is the set of indices compatible with everything  $x$  believes in  $i$

- Main issue:
  - No predictions about how strongly *believe* commits the agent to the prejacent.
  - *Dox* is defined as the set of indices compatible with everything the agent “believes” at the relevant index, so the issue of strength is pushed into the metalanguage.

## Two views on the strength of *believe*

- Gradability
- Subjective vs. objective epistemic modality
- Weak Believers
- Strong Subjective Believers

## Believe as a gradable predicate

- *Believe* can participate in comparative and equative constructions (if supported by gradable adverbs like *strongly*).
  - (6) a. He believes more strongly than I do that the organization of the executive branch of the federal government matters a great deal.  
b. Each [farmer] believes as strongly as the other that his crops will not survive another week without water [...].
  - (7) a. Atticus partially believes that prejudice exists because people do not understand each other [...].  
b. I strongly believe that life is too short to eat mediocre meals.  
c. This has taken me lots of research to come to this conclusion, but I believe 95 percent that it is.



## A gradable semantics for *believe*

- *Believe* encodes a measure function and requires that the agent's certainty in the prejacent meets some norm.

$$(8) \quad \llbracket \text{believe} \rrbracket^{w,t} = \lambda p \lambda d \lambda x . \mu_{x,w,t}(p) \geq d,$$

where  $\mu$  is a measure function parameterized by agents, worlds, and times

- Compositional details (Koev 2019)
  - In the absence of overt degree morphology, the degree argument is supplied by a covert POS morpheme (Cresswell 1976; Kennedy & McNally 2005).
  - Otherwise, the degree argument is filled by a degree modifier, e.g. *partially*.

## Subjective vs. objective epistemic modality

- Lyons 1977 (see also Kratzer 1981; Nuyts 2001; Papafragou 2006; Portner 2009):
  - **Subjective modality** merely describes the mental state of the agent. Expresses an opinion and need not be based on evidence. May be used to reduce the degree of public commitment to the prejacent.
  - **Objective modality** is based on evidence. Contributes to the propositional content of the sentence.
- Example
  - (9) Alfred must be unmarried. (Lyons 1977: 791-792)
    - a. Subjective: I (confidently) infer that Alfred is unmarried.
    - b. Objective: In the light of what is known, it is necessarily the case that Alfred is unmarried.

## More on subjectivity

- Subjectivity and strength are unrelated.
  - The subjectivity/objectivity distinction is about modal content and is orthogonal to the implied degree of strength, which is about modal force.
  - A subjective use does not entail a low degree of certainty in the prejacent.
  - Two different modal scales:  $\langle \textit{subjective}, \textit{objective} \rangle$  vs.  $\langle \textit{weak}, \dots, \textit{strong} \rangle$
- Lexicalization possible.
  - Items may lexicalize a particular flavor of epistemic modality.
  - Mental state predicates like *believe*, *think* or *doubt* are inherently subjective.
  - *Sure* is inherently objective (or at least can be read objectively).

# Weak Believers

- *Believe* invokes a degree of certainty that exceeds some contextually determined threshold.
- This threshold is typically 50% but can shift somewhat when there are several alternatives to the prejacent.
- We call this view **Weak Believers** (WB).

## Weak Believers (Hawthorne et al. 2016)

$\llbracket \text{POS} \rrbracket^{C,w,t}(\llbracket \text{believe} \rrbracket^{C,w,t}(p))(x)$  iff  $\mu_{x,w,t}(p) > \theta_{bel}$ ,  
where typically  $\theta_{bel} = 0.5$

## Strong Subjective Believers

- *Believe* conveys maximal subjective certainty.
- It differs from epistemic modals like *sure*, which we suggest encode high but objective certainty.
- Formally, we split the generic probability function  $\mu$  into two separate functions:  $Cr$  measures “credences” or subjective certainty while  $Pr$  measures objective certainty.
- We call this view **Strong Subjective Believers** (SSB).

### Strong Subjective Believers (Koev 2019)

$$\llbracket \text{POS} \rrbracket^{C,w,t}(\llbracket \text{believe} \rrbracket^{C,w,t}(p))(x) \quad \text{iff} \quad Cr_{x,w,t}(p) = 1$$

# Empirical evidence for each view

- WB
  - Neg-raising
  - Modal gradation
  - Hedging
- SSB
  - Missing uncertainty implicatures
  - Closure under conjunction
  - Scale structure

## Argument #1 for WB: Neg-raising

- *Believe* is a classic neg-raising predicate: *x doesn't believe p* often comes to mean *x believes not p*.
- Hawthorne et al. 2016 hypothesize a strict link between neg-raising and (modal) strength.

### The neg-raising/modal strength hypothesis

Neg-raising is licensed by with weak/non-strong modals (*think, want, ...*) but not by strong modals (*know, need, ...*).

- Since *believe* licenses neg-raising, it must have a weak/non-strong semantics.
- Caveat: Horn 1989 cites crosslinguistic evidence showing that strong modals occasionally license neg-raising.
- So neg-raising does not entail modal weakness.

## Argument #2 for WB: Gradation sequences

- If *believe* is weak, it should be possible to strengthen it by using a stronger modal, such as *know*. This seems possible.  

(10) Scientists believe there is water on Mars. In fact, they know it.
- However: *Know* need not carry a stronger modal force than *believe*.
- *Know* is “stronger” than *believe* in at least two other respects: factivity and evidence requirement.
- So (10) may simply involve a raising of the commitment of the speaker towards the prejacent, not that of the belief agent.



## Argument #3 for WB: Hedging

- The use of *believe* as a hedge to one's discourse commitments is often presented as an argument for WB.

(11) I believe it's raining, but I'm not sure it is.

- Notice: SSB is also able to account for such data, but it involves additional assumptions, i.e. the distinction between subjective and objective epistemic modality.

## Argument #1 for SSB: Missing uncertainty implicatures

- Belief attributions systematically lack “uncertainty” quantity implicatures.
  - (12) a. Kamala believes that America needs universal health care.
  - b. ↗ Kamala is not certain that America needs universal health care.
- This is as expected if *believe* conveys full (subjective) certainty.

## Argument #2 for SSB: Closure under conjunction

- A modal  $M$  is **closed under conjunction** iff  $M(p) \wedge M(q) \models M(p \wedge q)$ .
- Strong modals are closed under conjunction while non-strong modals are not.
  - (13)
    - a. It's certain that Sean is in Rome and it's certain that he is catholic.
    - b.  $\models$  It's certain that Sean is in Rome and that he is catholic.
  - (14) *Each week Jack spends (in no particular order) 3 nights at the local pub and gets drunk, 2 nights at the same pub but stays sober, and 2 nights at home where he also gets drunk. On a given night, I say:*
    - a. Jack is probably at the pub. T (chance =  $\frac{5}{7}$ )
    - b. Jack is probably drunk. T (chance =  $\frac{5}{7}$ )
    - c. Jack is probably at the pub drunk. F (chance =  $\frac{3}{7}$ )

## Argument #2 for SSB: Closure under conjunction

- *Believe* is closed under conjunction, like *certain* and unlike *probably*.
  - (15) a. Ron believes Mia is hawt and he also believes she is going to marry him.
  - b.  $\models$  Ron believes that Mia is hawt and that she is going to marry him.
- *Believe* exhibits the logical properties of a strong modal.
- Formally, SSB but not WB derives the closure property.

## Argument #3 for SSB: Scale structure

- Unger 1971
  - Gradable predicates can be **relative** or **absolute**.
  - In positive forms, relative predicates (*tall*) take vague standards of comparison while absolute predicates (*empty*, *full*) take scale endpoints.
- Kennedy and McNally 2005
  - The relative/absolute distinction is about scale structure.
  - Relative adjectives are associated with an open scale, so they cannot be modified by minimality, maximality, or proportional adverbs (cf. *\*slightly tall*, *\*completely tall*, *\*half tall*).
  - Absolute adjectives have scales with endpoints, and thus accept minimality, maximality, or proportional modifiers (cf. *slightly empty*, *completely empty*, *half empty*).
- *Believe* sentences
  - Accept minimality, maximality, and proportional modifiers (7).
  - *Believe* is associated with a totally closed scale, i.e.  $[0, 1]$ .
  - It must pick a scale endpoint (0 or 1) as a standard of comparison, just like SSB predicts.

# WB and SSB on hedging

- Divergent predictions
- Grammatical form
- Discourse structure

## Divergent predictions about hedging

- How do WB and SSB account for the hedging data?
  - (16) I believe the Giants will win, but I'm not sure they will.

$$\underline{\text{WB}}: \mu_{x,w,t}(p) > \theta_{bel} \wedge \mu_{x,w,t}(p) < 1 \quad \Delta p \wedge \neg \square p$$

$$\underline{\text{SSB}}: Cr_{x,w,t}(p) = 1 \wedge Pr_{x,w,t}(p) < 1 \quad \square_s p \wedge \neg \square_o p$$

Abbreviations:  $p = \llbracket \text{the Giants will win} \rrbracket$ ,  $\Delta =$  it is likely,  $\square =$  it is certain,  $s =$  subjective,  $o =$  objective.

- Each view takes the hedging data in stride, but WB wins out on simplicity.
- However: WB and SSB diverge in at least two further predictions that we argue favor SSB.

# WB on hedging sentences

- Grammatical form
  - WB draws a purely logical contrast in strength and predicts felicity across the board.
  - No difference whether (16) occurs in the **first** or **third person**, in the **present** or **past tense**, **unembedded** or **embedded** under hypotheticals like *suppose*.  
(We focus on these three factors because they may affect epistemic contradictions, e.g. Moore-paradoxical sentences.)
- Discourse status
  - (16) should be fine regardless of whether what is at-issue is the **prejacent** (i.e.  $QUD_1 = \textit{Will the Giants win?}$ ) or the **belief component** (e.g.  $QUD_2 = \textit{Why did you bet on the Giants?}$ ).
  - $QUD_1$ : Both clauses qualify the likelihood of the prejacent, so they address the same question.
  - $QUD_2$ : Both clauses convey the agent's attitude toward the prejacent.



# SSB on hedging sentences

- Grammatical form
  - SSB draws a contrast between stated private beliefs and incurred discourse commitments, and this requires a salient speech context.
  - So first person/present tense/unembedded forms – as most closely tied to the utterance context – are expected to be more natural than combinations of third person/past tense/embedded forms, which are further detached from the utterance context.
- Discourse factors
  - QUD<sub>1</sub>: The discourse topic invokes objective probabilities, so the speaker has to hedge in order to avoid too strong a commitment. This predicts felicity.
  - QUD<sub>2</sub>: The questions is asking about a personal estimate while the speaker additionally invokes objective certainty in the follow-up clause. This pivots to a different QUD and is likely to feel irrelevant.

## Summary of predictions

	1P+Pres+Main	3P+Past+Emb
WB	✓	✓
SSB	✓	?

Table: Grammatical form

	Prejacent	Belief
WB	✓	✓
SSB	✓	?

Table: Discourse factors

# Experimental evidence

- Experiment 1 (grammatical factors)
- Experiment 2 (discourse factors)
- General discussion

# Experiment 1: Grammatical factors

- Goal: Test the effects of grammatical factors on the acceptability of sentences in which *believe* is used as a hedge.
- Design:  $2 \times 2 \times 2$ , with the following factors:
  - Person: first vs. third
  - Tense: present vs. past
  - Position: main clause vs. embedded under *suppose*

# Experiment 1: Participants

- 96 participants from MTurk
- Residents in US (with US IP)
- Over 18 years of age
- 7 participants failed the exclusion criteria
- Final data set had 89 participants

# Experiment 1: Procedure

- Participants were not told the purpose of the study.
- Participants were explicitly directed not to judge based on real-world plausibility but on naturalness of sentences.

I believe Friends is among the top ten American series, but I'm not sure.

How natural does this sentence sound to you?

very unnatural very natural

Further comments (if any):

NEXT

Figure: Sample trial in Experiment 1.

# Experiment 1: Materials

- 16 sentences  $\times$  8 conditions = 128 test sentences
- 18 filler sentences, 9 good and 9 bad ones
- All sentences of the form of  $believe_x p \wedge \neg sure_x p$
- Fillers contained different predicates, e.g. *guess*, *know*, *imagine*
- Exclusion criterion: more than 3 bad fillers rated higher than 40%

## Experiment 1: Sample items

- (17) I believe it's raining, but I'm not sure. (target sentence)
- (18) I guess she's from Canada, but I'm not sure. (good filler)
- (19) I know what you are saying is true, but it's false.(bad filler)



# Experiment 1: Design

- 8 lists with 34 final target sentences (Latin square design)
- 2 sentences per condition
- Item order was pseudo-randomized
- Each list started with 1 good and 1 bad filler
- Experiment lasted on average 8 minutes

## Experiment 1: Data summary

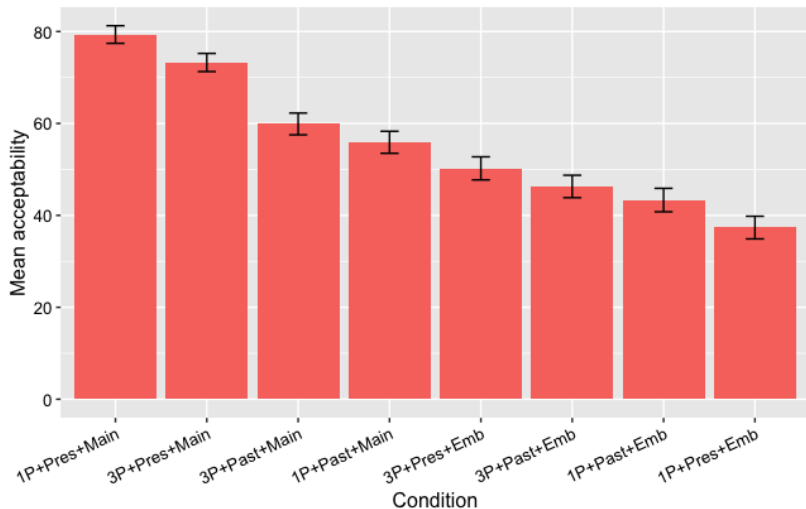


Figure: Mean values of responses per condition. Error bars indicate SE.

# Experiment 1: Data analysis

- Fitted a mixed effects linear regression model following Barr et al. (2013)'s **best path algorithm**.
  - Fixed effects: Position, Tense, Person, and all their interactions.
  - Random effects: Intercepts for participant and item. By-participant slopes for Position.
- Tested for the significance of the main effects and the various interactions through model comparison.

# Experiment 1: Data analysis

- Significant difference in main effect
  - Position ( $\chi^2(1) = 69, p < 0.001$ ): Main > Embedded
  - Tense ( $\chi^2(1) = 55, p < 0.001$ ): Present > Past
  - Person ( $\chi^2(1) = 8, p < 0.01$ ): 3rd > 1st
- Significant difference in interactions of
  - Position and Tense ( $\chi^2(1) = 66, p < 0.001$ )
  - Position and Person ( $\chi^2(1) = 13, p < 0.001$ )
  - Position, Tense, and Person ( $\chi^2(1) = 16, p < 0.001$ )
- No significant effect in
  - Tense and Person ( $\chi^2(1) = 0.02, p = 0.88$ )

# Experiment 1: Data summary

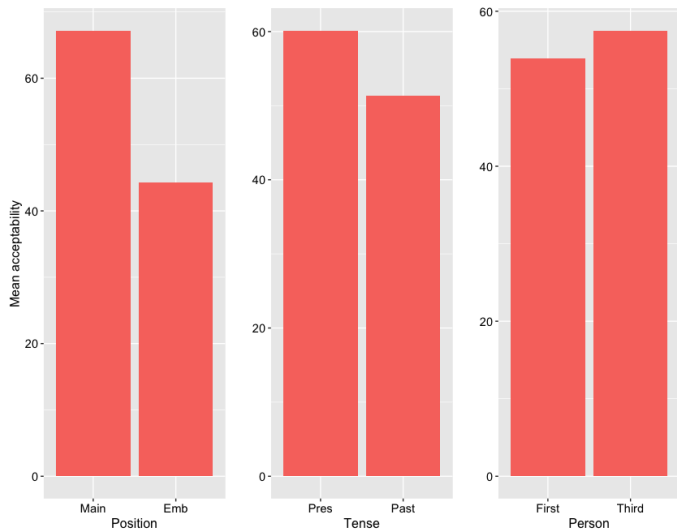


Figure: Mean values of responses per fixed factor.

# Experiment 1: Results

- As predicted by SSB, participants rated present tense/main clause forms higher than past tense/embedded clause forms.
- Contrary to SSB prediction, participants rated third person forms higher than first person forms.

## Experiment 2: Discourse factors

- Goal: Test the acceptability of sentences in contexts where the prejacent was at-issue to those where the belief component itself was at-issue.
- A single predictor At-issueness with two levels: Prejacent vs. Belief component.

## Experiment 2: Participants

- 62 participants from MTurk
- Residents in US (with US IP)
- Over 18 years of age
- 14 participants failed the exclusion criteria
- Final data set had 48 participants



## Experiment 2: Procedure

- The same procedure as Experiment 1

## Experiment 2: Materials

- 4 sentences  $\times$  2 conditions = 8 test items
- 6 filler sentences, 3 good and 3 bad ones
- All items were question/answer pairs, with answers of the form of  $believe_x p \wedge \neg sure_x p$
- Fillers items contained polar or constituent questions
- Fillers contained different predicates, e.g. *guess*, *know*, *imagine*
- Exclusion criterion: All 3 bad fillers rated higher than 50%

## Experiment 2: Sample items

### Prejacent condition

- (20) A: Are Nike Zoom the best running shoes?  
B: I believe they are, but I'm not sure.

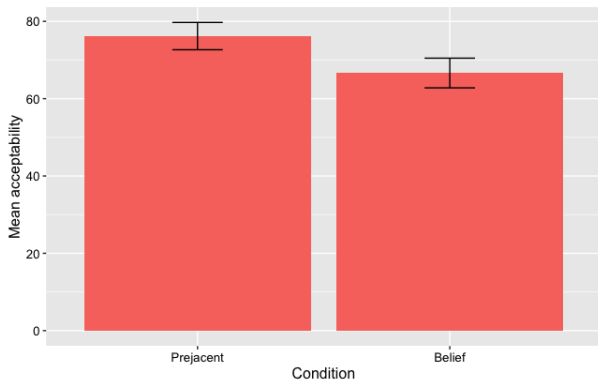
### Belief condition

- (21) A: Why did you buy Nike Zoom running shoes?  
B: I believe they are the best, but I'm not sure.

## Experiment 2: Design

- 2 lists with 10 question/answer pairs (Latin square design)
- 2 sentences per condition
- Item order was pseudo-randomized
- Each lists started with 1 good and 1 bad filler
- Experiment lasted in average about 4 minutes

## Experiment 2: Data summary



**Figure:** Mean values of responses for At-issueness condition. Error bars indicate standard error.

## Experiment 2: Data analysis

- We fitted a mixed-effect linear regression model to the data.
- A model comparison found a significant effect for At-issueness ( $\chi^2(1) = 11, p < 0.001$ ).
- In line with SSB's predictions, the prejacent condition sentences rated as more acceptable than Belief condition sentences.

# General discussion: Experiment 1

## Findings

- Present > Past
- Main > Embedded
- Third Person > First Person

## WB

- None of these contrasts are predicted.

## SSB

- Correctly predicts the first two contrasts in Tense and Embedding: Past and embedded belief reports are further detached from the utterance context.
- But why are 3P forms more acceptable than 1P forms?

## General discussion: Experiment 1

- Suggestion: 3P belief reports naturally evoke a secondary speech context, as people have no direct access to other people's mental states.
- So *John believes it's raining* will typically be taken to imply *John said it's raining*, where  $Cr_{x,w,t}$  is such that  $x = \text{John}$ ,  $w = \text{the world of John's utterance}$ ,  $t = \text{the time of John's utterance}$ .
- The subjective/objective contrast is “shifted” to a secondary context.
- But why were 3P forms rated as more acceptable than 1P forms? Because 1P forms cannot undergo a shift to a secondary speech context: 1P belief reports are based on direct experience.



# General discussion: Experiment 2

## Findings

- Prejacent > Belief component

## WB

- Predicts no contrast.
- Perhaps the contribution of epistemic modals is generally more difficult to be construed as at-issue than the prejacent (cf. Papafragou 2006; but see Simons 2007)?
- However: SSB independently predicts this contrast and thus has more predictive power.

## SSB

- Correctly predicts the contrast because in the Belief condition there is a topic change.

- The bigger question: How far does the claimed distinction between subjective vs. objective certainty cut into the epistemic domain?
- Hypothesis: All epistemic modals are lexically specified as subjective (*believe, doubt*), objective (*sure*), or either (*might, must, possible, likely, certain*).
- Modals from the latter, “neutral” category are not lexically ambiguous. They are underspecified and their interpretation is context dependent.

Thank you!