Disagreements do not automatically raise the standard of precision

Speakers often choose to utter imprecise sentences that strictly speaking are false (1a). The *standard of precision* [1-3] governing a discourse can be negotiated through metalinguistic disagreements: in (1b), Andy's challenge signals that a stricter standard of precision (SoP) should be adopted. Here we investigate whether metalinguistic disagreements like (1b) result in an automatic update of the SoP. We consider two hypotheses: **Hypothesis 1 (H1)** states that challenging the SoP automatically updates this discourse parameter, superseding previous parametrizations. **Hypothesis 2 (H2)** states that metalinguistic challenges act as a request to shift the SoP, but do not directly update it. Unlike H2, H1 predicts that disagreements should decrease the acceptability of a previous imprecise utterance. Contra H1, we find that imprecise utterances continue to be perceived as felicitous even after the SoP has been challenged, suggesting that any potential updates to the SoP ought to take place in subsequent conversational moves.

Experiment 1 (Exp1): We created twenty-four five-point scales instantiating different Maximum Standard adjectival properties (e.g., *empty*) to varying degrees (Fig.1a). Each scale was normed (n=30) to ensure that the lower scalepoints (1-4) tolerated some amount of imprecision. The goal of Exp1 (n=30) was to gather interpretational preferences for individual scale points in isolation to be used as a baseline in the analysis of Experiment 2. Participants saw individual images accompanied by a description of the form '*This* [object] *is* [adjective]' (Fig. 2a), and were instructed to choose one of three answers: '*Yes*', '*Unsure*' or '*No*'. Exp1 results are shown in Fig. 2b.

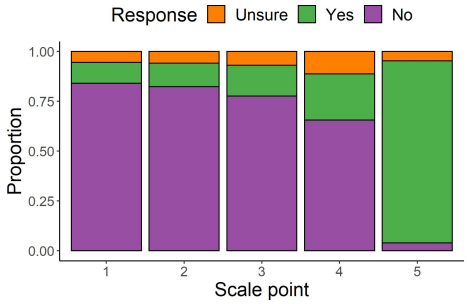
Experiment 2 (Exp2): (n=60) The goal of Exp2 was to assess whether metalinguistic disagreements modulate the acceptability of imprecise utterances. Participants saw the same stimuli used in Exp1 with the only difference that the initial assertion '*This* [object] *is* [adjective]' was followed by an utterance of the form '*No, this* [object] *is not* [adjective]' (Fig. 3a). Participants' task was to choose one of three options: *'Both of them can be right,' 'Only the* {*first, second*} *speaker is right*. The scale points in the 24 scales tested were distributed in 5 lists following a Latin-square design. Twenty-four disagreements about properties not subject to imprecision (e.g., *checkered*) were included as fillers.

Results: Responses (Fig. 3b) were binarized such that selections of '*both of them can be right*' (henceforth *Both*) were coded as 1, while the remaining two levels were coded as 0. A logistic mixed effects regression model was fitted to this new binomial variable using SCALE POINT as a fixed effect. Scale point 5 (S5) was coded as the reference level. Random intercepts and slopes by items and participants were also included. All comparisons were significant, with lower scale points (S1-S4) receiving higher proportions of *Both* responses compared to S5 (all p's < 0.05). Next, we constructed two new binary variables coding whether participants selected '*Only the* {*first, second*} *speaker is right*' (henceforth *First* and *Second*) respectively. The same procedure was followed for Exp1 '*Yes*' and '*No*' responses. The four binomial variables were appended and coded based on 1) whether the observation belonged to Exp1-2 (EXPERIMENT); and 2) whether the imprecise utterance was *accepted* (i.e., '*Yes*' in Exp1, and '*First*' in Exp2) or *rejected* (i.e., '*No*' in Exp1, and '*Second*' in Exp2, see Fig. 4). We refer to this factor as ACCEPTABILITY. A series of mixed effects models were fitted to the data pertaining to each scalepoint, with EXPERIMENT, ACCEPTABILITY and their interaction as fixed effects. Random intercepts and slopes by item and participant were also included. The interactions were significant in S1-4 (all p's < 0.05; S5: p's > 0.05). Simple effect analyses revealed the interactions were driven by higher rates of '*No*' responses compared to *Second* responses (S1-4: all p's < 0.05). No significant differences were detected between '*Yes*' and *First* responses (S1-5: all p's > 0.05).

Discussion & Conclusion. Our results suggest that imprecise utterances are not deemed unacceptable when embedded in a disagreement dialogue. This is shown by the fact that *First* responses were comparable to '*Yes*' responses in S1-4. Conversely, proportions of *Second*—a choice compatible *only* with a higher SoP—were lower than '*No*' responses in S1-4. These lower rates were due to participants displaying a higher preference for *Both*—an option compatible with a lower SoP—in S1-4 compared to S5. The current findings are therefore incompatible with H1, but can be better accommodated by H2. In further research, we address how the discourse commitments [4] incurred by subsequent conversational moves (e.g., concessions, vs. retractions) update the SoP.

(1) a. *Shelly:* This bottle is empty.
b. *Andy:* No, this bottle is not empty, there's a bit of water in it.

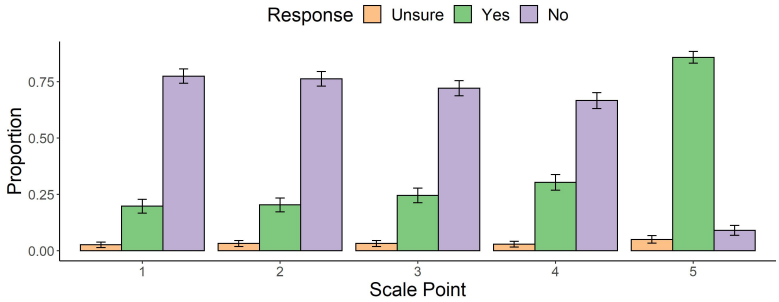
(a) Norming Study Item Example.



(b) Norming Results.

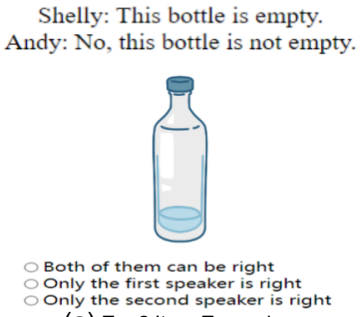
Figure 1: Norming Study.

(a) Exp1 Item Example.

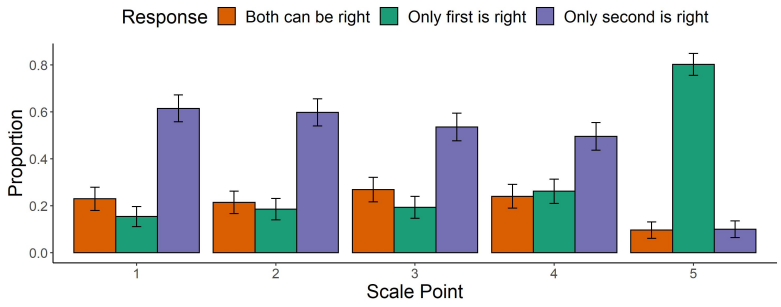


(b) Exp1 Results.

Figure 2: Experiment 1.

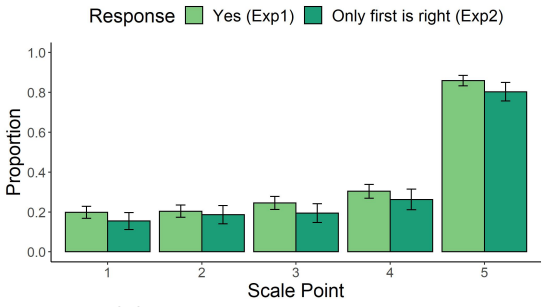


(a) Exp2 Item Example.

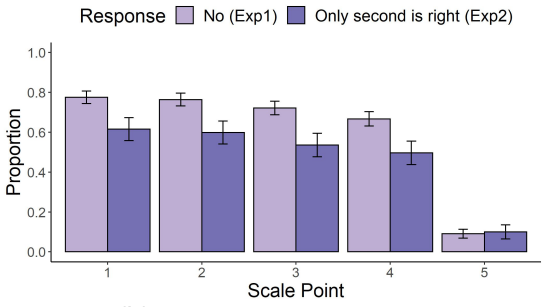


(b) Exp2 Results.

Figure 3: Experiment 2.



(a) Imprecise utterance accepted.



(b) Imprecise utterance rejected.

Figure 4: Exp1-2 comparison.

References: [1] Lewis, D. (1979). *Scorekeeping in a language game.* J. Philos. Log. | [2] Lasersohn, P. (1999). *Pragmatic halos* Lang. | [3] Klecha, P. (2018). *On unidirectionality in precisification*. L&P. | [4] Lauer, S. (2012). *On the pragmatics of pragmatic slack*. Proc. SuB.