Meaning adaptation in the discourse dynamics of imprecision

Yifan Wu (yw2578@cornell.edu)
Department of Linguistics, Cornell University

Helena Aparicio (haparicio@cornell.edu)
Department of Linguistics, Cornell University

Abstract

Speakers often communicate imprecisely, using expressions that are strictly speaking false yet felicitous. The degree to which imprecision is tolerated in a discourse is governed by the Standard of Precision (SoP). While it is known that contextual factors can modulate the SoP (e.g., Beltrama & Schwarz, 2024; Mathis & Papafragou, 2022), less is known about the discourse dynamics of imprecision, particularly how speakers coordinate on the SoP during conversation. Previous accounts have claimed that implicit negotiations of the SoP are unidirectional, namely they only work upward, not downward (Klecha, 2018; Lewis, 1979). Here, we investigate whether parallel asymmetries arise when comprehenders adapt their SoPs in response to an interlocutor's precision preferences. We present results from two first-person dialogue experiments, where participants interacted with an interlocutor that displayed either a preference for precision or imprecision. Our results show bidirectional SoP adaptation effects: exposure to lower standards increased tolerance for imprecision, while exposure to higher standards reinforced stricter thresholds. These updates persisted beyond the dialogue, suggesting that exposure to (im)precise speakers modulates not only interpretations within a discourse, but also beyond the conversation. More broadly, our study provides a novel framework for studying real-time dynamic meaning negotiations during conversation.

Keywords: Imprecision, standard of precision, first-person dialogue experiments, discourse processing, semantic and pragmatic adaptation

Introduction

In everyday communication, speakers often use linguistic expressions imprecisely, even though the propositional content of an imprecise utterance is, strictly speaking, false (Aparicio, Xiang, & Kennedy, 2015; Beltrama & Schwarz, 2024; Burnett, 2014; Kao, Wu, Bergen, & Goodman, 2014; Klecha, 2018; Krifka, 2002, 2007; Lasersohn, 1999; Leffel, Xiang, & Kennedy, 2016; Lewis, 1979; Ronderos, Noveck, & Falkum, 2024; Syrett, Kennedy, & Lidz, 2010; Wu & Aparicio, 2025; a.o.). This phenomenon is exemplified in (1), where Speaker 1 describes the bottle in Figure 1 as '*empty*', despite being aware that it contains some water. Slightly more formally, imprecision occurs when an utterance is perceived as pragmatically felicitous, even though its communicated content is weaker than its literal interpretation. In this respect, imprecision differs from other better-studied pragmatic phenomena, such as scalar implicatures, in which pragmatic reasoning serves to strengthen, rather than weaken, the literal meaning (Lauer, 2012, 2013).

Imprecision is a cross-categorical phenomenon that extends beyond maximum standard adjectives such as *empty*. It can also affect lexical categories, such as incremental theme verbs (e.g., *eat an apple*), prepositions (e.g., *in*), and round numerals (e.g., *one hundred*), among others. Previous studies have identified several factors that determine the acceptability of an imprecise utterance. Ronderos et al. (2024) show that the tolerance for imprecision gradually decreases as the imprecise interpretation deviates from the utterance's precise—i.e., literal—meaning (e.g., the maximal degree of emptiness in Figure 1, Cf. Kennedy, 2007; Kennedy & McNally, 2005; Rotstein & Winter, 2004). Previous work has demonstrated that the discourse context strongly influences the felicity of an imprecise utterance, with low-stakes contexts, where minor deviations from truth are inconsequential, generally allowing for more imprecision, compared to high-stakes situations (Aparicio Terrasa, 2017; Burnett, 2014; van Der Henst, Carles, & Sperber, 2002; Mathis & Papafragou, 2022; Ronderos et al., 2024; Solt, 2015; Solt, Cummins, & Palmović, 2017; a.o.). Finally, recent studies show that comprehenders incorporate social information about the speaker to guide their inferences about imprecision, associating *chill* speakers with a greater tendency to speak imprecisely (Beltrama & Schwarz, 2021, 2022, 2024; Beltrama, Solt, & Burnett, 2023).

(1) a. *Speaker 1:* This bottle is empty.
b. *Speaker 2:* I disagree. This bottle is not empty.

(2) a. *Speaker 1:* This bottle is not empty.
b. *Speaker 2:* I disagree. This bottle is empty.

Figure 1: Metalinguistic disagreements.

However, much less is understood about the *discourse dynamics* of (im)precision, particularly how conversational participants *coordinate* on the degree of imprecision that is acceptable within a given discourse. Here, we adopt the proposal that the degree of imprecision tolerated in a conversation is governed by the *Standard of Precision* (SoP), which we model as a latent discourse parameter (Klecha, 2018; Lasersohn, 1999; Lauer, 2012; Lewis, 1979; a.o.). Because the SoP is not observable, conversational agents must rely on contextual cues (some of which have already been mentioned above) to infer likely parametrizations of the SoP. While in most cases speakers and listeners successfully align on the value of the SoP, this implicit coordination process is not infallible; occasionally, interlocutors assume conflicting parametrizations that can eventually cause conversational disruptions. When such misalignments occur interlocutors can implicitly negotiate the parametrization of the SoP through metalinguistic disagreements (Barker, 2002, 2013; Horn, 1989), where one interlocutor challenges the assertability of a previous utterance, as seen in Figure 1. Lewis (1979) originally claimed that such metalinguistic challenges are asymmetric: while a metalinguistic challenge can be effectively used to raise or strengthen the SoP—that is, it can serve to precisify—it rarely succeeds in lowering or weakening it. For example, in Figure 1, an imprecise utterance like (1a) can be contested by (1b), signaling that a stricter standard should be adopted. In contrast, attempts to lower the SoP in order to increase the tolerance for imprecision, as in (2b), have been claimed to be infelicitous and ineffective (Cf. Klecha, 2018).

This proposed unidirectionality raises the question of whether similar asymmetries extend to how much interlocutors are willing to adapt to the degree of (im)precision displayed by their conversational partner. While semantic and pragmatic adaptation has been previously studied in domains such as vague quantifiers, uncertainty expressions and gradable adjectives (Heim, Peiseler, & Bekemeier, 2020; Pecsok & Aparicio, 2024; Schuster & Degen, 2019, 2020; Xiang, Kramer, & Kennedy, 2020; Yildirim, Degen, Tanenhaus, & Jaeger, 2016; a.o.), little is known about how adaptable (im)precise meanings are.

Here, we experimentally investigate the discourse dynamics of imprecision using what we refer to as *first-person* dialogues, where experimental participants actively engage in the communicative exchange. We hypothesize that a task in which participants assume the role of interlocutors will offer a more sensitive method for detecting how speakers accommodate to varying precision standards in real-time interaction. In this respect, the current work differs from previous adaptation studies, which either did not involve dialogues or positioned participants as bystanders to the communicative exchange, rather than active interlocutors. In contrast, our experimental design assigns participants the role of respondents, requiring them to take the second turn in a dialogue by responding (agreeing or disagreeing) to a purported interlocutor displaying a particular SoP preference: Experiment 1 tests exposure to an imprecise interlocutor, while Experiment 2 examines responses to a precise one.

We consider two hypotheses: 1) Our first hypothesis (H1) proposes that adaptation effects will be asymmetric, reflecting claims in the literature on metalinguistic disagreements that such disagreements can strengthen, but not weaken, the SoP (Klecha, 2018; Lewis, 1979). Specifically, exposure to a precise speaker should cause participants to raise their SoP, whereas exposure to an imprecise speaker should not lead to a lowering of their SoP. H1 therefore predicts strengthening effects in Experiment 2, but no adaptation effects in Experiment 1; 2) Hypothesis 2 (H2) proposes that adaptation effects should be bidirectional: comprehenders should loosen their baseline SoP when exposed to an imprecise speaker, and strengthen it when exposed to a precise speaker. H2 therefore predicts adaptation effects in both experiments. Our results strongly support H2: exposure to both lower and higher standards in first-person dialogues effectively shifted participants' precision thresholds in the expected direction. Results also reveal that these updates persist once the discourse concludes. Additionally, our findings suggest that first-person dialogues provide an effective means of capturing the discourse dynamics of imprecision.

Experiment 1

The goal of Experiment 1 was to determine whether exposure to imprecise utterances led participants to lower their SoP in order to align with their interlocutor's imprecise standard.

Method

Materials, design and procedure We constructed five-point scales instantiating varying degrees of maximum standard absolute adjectival properties (e.g., *empty bottle*, Figure 2). Each scale was normed to ensure that lower scale-points (S1-S4) tolerated some degree of imprecision. Participants judged whether a given scale-point could be described using the relevant adjectival property (e.g., *empty*) by responding '*Yes*', '*No*' or '*Unsure*'. Results (Figure 3) confirm that while precise interpretations were preferred overall, there was a non-negligible tolerance for imprecision at all lower scale-points, gradually increasing as the scale approached the endpoint (green bars in Figure 3).

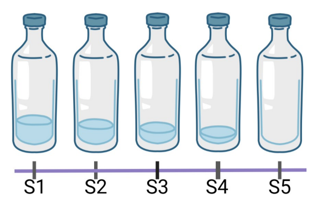


Figure 2: Scale-points.

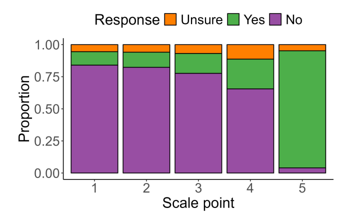
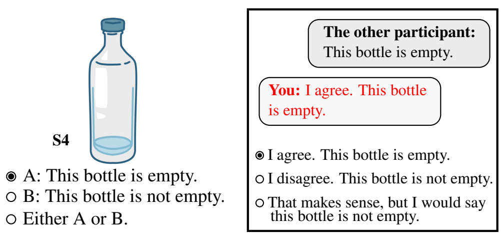


Figure 3: Norming results.

The experiment tested S4, the near-maximal scale-point that displayed the highest tolerance for imprecision in our norming study. We also included S5, which was compatible only with a precise interpretation of the predicate (e.g., *empty*), and S1, the scale-point that displayed the least tolerance for imprecision, as controls. Experiment 1 followed a three-block adaptation paradigm (Heim et al., 2020). In the pre-exposure phase, **Block1** (**B1**), participants provided interpretational preferences for each scale-point by selecting one of three options: '*A: This* [object] *is* [ADJECTIVE]', '*B: This* [object] *is not* [ADJECTIVE]', or '*Either A or B*' (Figure 4a). The goal of **Block2** (**B2**) was to determine whether exposure to an imprecise utterance caused participants to adapt towards a lower SoP than that displayed at baseline (B1). Participants engaged in a conversation with a purported interlocutor to describe the same object they had seen in B1. Participants were informed that they would be remotely paired with another participant (*the other participant*) who could see the same object on their screen. In reality, this interlocutor was a chatbot (henceforth, *bot*) designed to simulate real-time conversational exchanges with a specific SoP preference.

Experiment 1 featured an *imprecise bot* that initiated the dialogue with the experimental participant. The imprecise bot accepted the predication (i.e., '*This* [OBJECT] *is* [ADJECTIVE]') at S4 and S5, and rejected it at S1 ('*This* [OBJECT] *is not* [ADJECTIVE]'). Participants were asked to respond to the bot's initial utterance by choosing one of three options: '*I agree. This* [OBJECT] *is* [ADJECTIVE]', '*I disagree. This* [OBJECT] *is not* [ADJECTIVE]' or '*That makes sense, but I would say that this* [OBJECT] *is* (*not*) [ADJECTIVE]' (Figure 4b). At S4, our critical scale-point, agreement-responses indicate that participants accepted the low SoP assumed by the bot. A disagreement-response, i.e., a rejection of the imprecise stance, indicates that participants adopted a higher SoP compared to their interlocutor's. Finally, the third type of response consists of a nuanced form of disagreement where participants acknowledge the bot's perspective while maintaining their own stricter SoP. Regarding the control S1, we expected imprecise interpretations to be less acceptable in S1 compared to S4, as S1 is further from the maximal endpoint (S5) than S4. Since the bot rejected the imprecise predication at S1, we expected that participants would be willing to update their standard upon exposure to the bot's utterance, potentially rejecting the predication at higher rates in B2 (i.e., *agree*-responses) compared to B1 (i.e., *B*-responses). S5 functioned as a truth-conditional control that did not involve any meaning uncertainty about whether the predicate applied to the object or not. We therefore expected that participants would largely accept the predication in B1 and would align with the bot's acceptance of the predication in B2. We therefore did not anticipate any changes between B1 and B2.



(a) Block 1/3. (b) Block 2.

Figure 4: Experiment 1 item example featuring an imprecise participant who accepts S4 in Blocks 1 and 3 and maintains their imprecise SoP in response to the bot in Block 2.

Finally, **Block 3** (**B3**), the post-exposure block, was an exact replication of B1 (Figure 4a). This block sought to capture whether the potential changes observed in B2 persisted beyond the dialogue. Participants gave responses for each block within trials, i.e., three responses per trial. Item conditions were distributed across three lists following a Latin Square design. This ensured that every participant only saw one scale-point per scale. Each list also included 24 filler items featuring color adjectives. All visual displays included stereotypical shades of the relevant color (Figure 5). Filler trials were later used as attention checks to ensure that participants stayed focused on the task. Fillers followed the same structure as experimental items. The order of the trials within each list was randomized and participants were randomly assigned to one of the three lists.

Figure 5: Filler item examples: Left: '*blue flag*'; Right: '*green circle*'.

The experiment was conducted remotely through the *PCIbex Farm* platform (Zehr & Schwarz, 2018). Participants began by providing informed consent, completing a demographic questionnaire, and completing three practice trials to familiarize themselves with the setup and response protocol before proceeding to the main experiment.

Participants Participants consisted of 60 native speakers of American English who were at least 18 years old and self-reported having normal vision with no colorblindness. All participants were recruited through the web platform *Prolific*. Participation was compensated at a rate of $15 per hour. Participants were excluded if their accuracy on attention check trials fell below 90%. One participant failed to meet the accuracy threshold and one participant was excluded due to data collection errors, resulting in a final sample of 58 participants.

Results

Results from Experiment 1 are visualized in Figure 6. To determine whether exposure to the bot's imprecise utterances caused participants to shift the SoP, we compared B1 and B2.

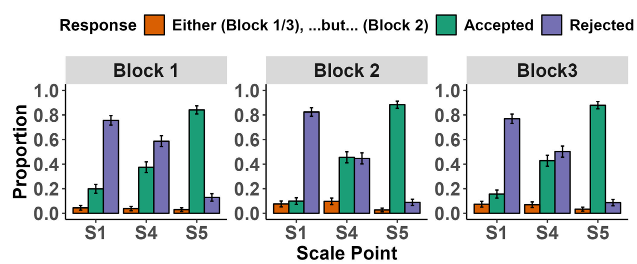
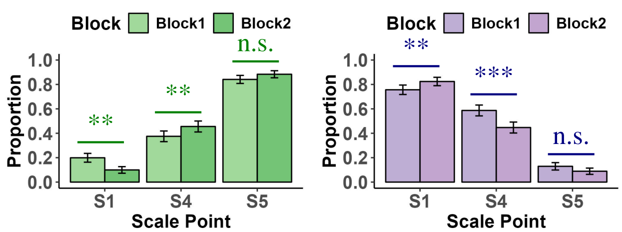


Figure 6: Experiment 1 results.

The critical scale-point S4 was consistently accepted by the bot. If exposure to lower standards influenced the participants' SoPs, we expected increased acceptance and decreased rejection of the imprecise utterance at S4, reflecting adaptation to the bot's standard during the dialogue. Therefore, we compared *A*-responses in B1 to agreement replies (henceforth *Agree*-responses) in B2, as these responses indicate that participants accepted the imprecise utterance and aligned with the lower SoP. Similarly, *B*-responses in B1 were compared to disagreement replies (henceforth *Disagree*-responses), which indicate participants' rejection of the imprecise utterance and alignment with the higher SoP. The '*Either A or B*' option in B1 was not directly comparable to the nuanced disagreement option ('*That makes sense, but I would say...*') in B2, as the former is compatible with both lower and higher SoPs, while the latter acknowledges the bot's conflicting standard and explicitly disagrees with it.

Responses were coded into two binary variables for analysis. In B2, *Agree*-responses were coded as 1, while all other responses were assigned 0, indicating acceptance of the imprecise utterance. A second variable was created where *Disagree*-responses were coded as 1 and all other responses as 0, reflecting rejection. The same coding was applied in B1, with *A*-responses coded as 1 in the acceptance variable (all other responses coded as 0), and *B*-responses coded as 1 in the rejection variable (with all other responses coded as 0). A mixed-effects logistic regression model was fit to each binary variable containing data from S4. The model included BLOCK (B2 vs. B1) as a fixed effect predictor, with B1 as the reference level. Each model included by-item and by-participant random intercept and slopes. Results showed significantly higher proportions of responses accepting the lower standard in B2 compared to B1 ($\beta = 0.73$, $z = 3.25$, $p < 0.01$, Figure 7a) and significantly lower rejection proportions in B2 compared to B1 ($\beta = -0.90$, $z = -3.67$, $p < 0.001$, Figure 7b).

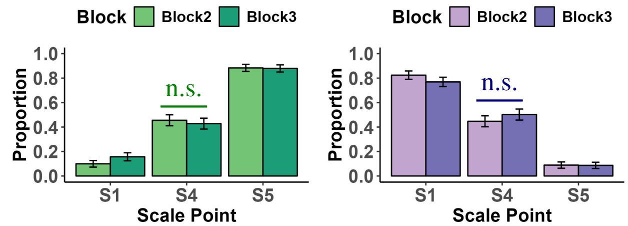


(a) Predication accepted. (b) Predication rejected.

Figure 7: Experiment 1: Block 1-2 comparison results.

In order to test whether the updates detected in B2 persisted post-dialogue in B3, we compared B2 and B3. The same coding system used for B1 was applied to the data pertaining to B3 (Figure 8), and the same model constructed for the comparison of B1 and B2 was fit to the data pertaining to B2 and B3 of S4, treating B2 as the reference level. Model results revealed no significant differences between B2 and B3 in either acceptance ($\beta = -0.21$, $z = -1.00$, $p > 0.1$, Figure 8a) or rejection rates ($\beta = 0.21$, $z = 0.79$, $p > 0.1$, Figure 8b).

To further confirm that the lack of differences between B2 and B3 reflected the persistence of the adaptation effects observed in B2, we directly compared post-dialogue B3 to the baseline B1 using the same binary coding logic and mixed-effects logistic regression models as before. Results revealed a significant increase in *A*-responses—responses accepting the imprecise predication—in B3 compared to B1 ($\beta = 0.54$, $z = 2.46$, $p < 0.05$). We also detected a significant decrease in *B*-responses—responses indicating rejection of the imprecise



(a) Predication accepted. (b) Predication rejected.

Figure 8: Experiment 1: Block 2-3 comparison results.

predication—in B2 compared to B1 ($\beta = 0.22$, $z = -3.32$, $p < 0.001$).

Finally, S1 and S5 behaved as expected (see Figure 7). At S1, exposure to the bot's rejection of the imprecise predication led to strengthening effects, evidenced by significantly lower acceptance and significantly higher rejection rates in B2 compared to B1. In contrast, no significant updates were observed for S5.[1]

Discussion

The comparison analyses between B1 and B2 at S4 reveal that acceptance rates of imprecise utterances in B2 (i.e., *Agree*-responses) were significantly higher than in B1 (i.e., *A*-responses), while rejection rates in B2 (i.e., *Disagree*-responses) were significantly lower than in B1 (i.e., *B*-responses). These findings indicate that exposure to the bot's consistent endorsement of S4 in first-person dialogues enhanced the acceptability of imprecise utterances, effectively causing participants to adapt towards a lower SoP. The current results are therefore not compatible with Hypothesis 1, which does not predict weakening effects on the SoP, but support Hypothesis 2, which posits bidirectional SoP modulation.

Finally, comparison analyses between B2 and B3 indicate that the adaptation effects observed in B2 for S4 persisted beyond the dialogue, as shown by the comparable acceptance rates of imprecise utterances in B2 and B3 (i.e., *Agree*-responses in B2 and *A*-responses in B3), and the comparable rejection rates in B2 and B3 (i.e., *Disagree*-responses in B2 and *B*-responses in B3). Further evidence that the adaptation effects carry over into B3 comes from the B1-B3 comparison, which revealed a significant increase in acceptance rates (i.e., *A*-responses) and a significant decrease in rejection rates (i.e., *B*-responses) in B3 compared to B1—a pattern that mirrors the results observed in the B1-B2 comparison. Taken together, these results suggests that adaptation to a lower standard in a first-person dialogue persists even after the dialogue concludes, potentially reflecting adjustments to individuals' subjective imprecision thresholds.

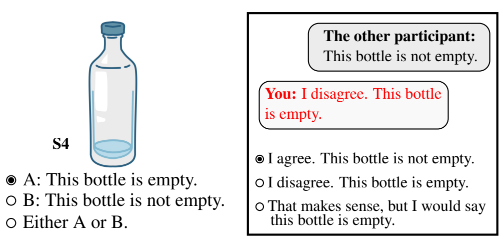
[1]Control-condition analyses were conducted using similar procedures to those outlined for S4, with modifications corresponding to the specific replies for each of the scale-points. Due to space limitations, we gloss over the details of the statistical analyses for the two control scale-points.

Experiment 2

Experiment 2 aimed to evaluate whether exposure to an interlocutor who always speaks precisely leads to a decrease in tolerance for imprecision, strengthening participants' SoPs.

Method

Materials, design and procedure Experiment 2 mirrored Experiment 1 with one crucial modification: the bot displayed a preference for *precision* by accepting the predication only at S5, and rejecting it both at S1 and S4. **Block 1** (**B1**) and **Block 3** (**B3**) were identical to Experiment 1 (see Figure 9a). As in Experiment 1, **Block 2** (**B2**) engaged participants in a first-person dialogue that was always initiated by *the other participant*, i.e., the bot. In Experiment 2, the bot displayed a preference for precision by consistently rejecting the predication at S1 and S4 (i.e., '*This* [OBJECT] *is not* [ADJECTIVE]') and accepting it at S5 (i.e., '*This* [OBJECT] *is* [ADJECTIVE]'). Participants were instructed to respond by selecting one of three options: '*I agree...*', '*I disagree...*', or '*That makes sense, but I would say...*' (Figure 9b). At the critical scale-point S4, an agreement response indicated alignment with the bot's precise SoP, while a disagreement response reflected a rejection of the precise SoP in favor of a more imprecise standard. The third option, nuanced disagreement, acknowledged the bot's perspective while maintaining the participants' original, lower SoP. The visual stimuli, filler trials, and procedure were the same as in Experiment 1.



(a) Block 1/3. (b) Block 2.

Figure 9: Experiment 2 item example, featuring an imprecise participant who accepts S4 in Blocks 1 and 3 and maintains their imprecise SoP in response to the bot in Block 2.

Participants The same recruitment criteria used in Experiment 1 were followed in Experiment 2, with 60 participants recruited initially. Three participants were excluded due to failure to meet the accuracy threshold in attention check trials, leaving a total of 57 participants.

Results

The results of Experiment 2 are illustrated in Figure 10. As in Experiment 1, S1 and S5 served as controls.[2] In Experiment 2, the bot consistently rejected S4, allowing us to investigate

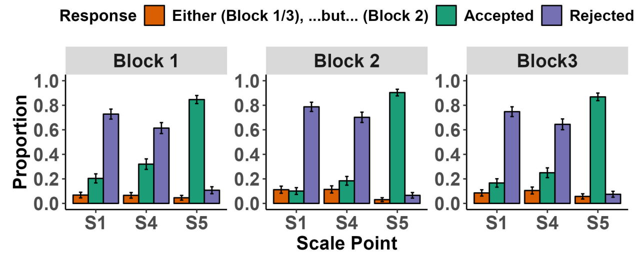
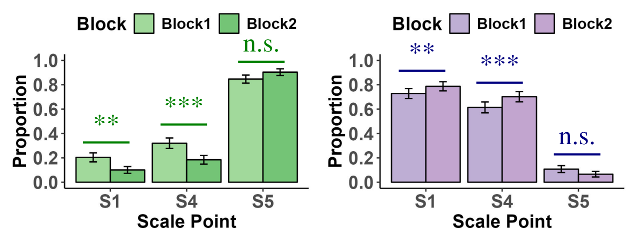


Figure 10: Experiment 2 results.

whether exposure to stricter standards would prompt participants to strengthen their SoP. To statistically assess this question, we compared the S4 data between B1 and B2 using the same coding scheme as in Experiment 1, but adjusted for the bot's precise preference. Thus, in Experiment 2, the independent variable representing acceptance of the imprecise predication coded *A*-responses in B1 and *Disagree*-responses in B2 as 1 (with all other responses coded as 0). Conversely, the independent variable representing rejection of the prediction coded *B*-responses in B1 and *Agree*-responses in B2 as 1 (with all other responses coded as 0). As in Experiment 1, we fit a mixed-effects logistic regression model to each independent variable separately, predicting the response variable from BLOCK. All models included by-item and by-participant random intercepts and slopes. The analyses confirmed significantly higher rejection rates in B2 compared to B1 ($\beta = 0.83$, $z = 3.64$, $p < 0.001$, Figure 11b), while acceptance rates decreased significantly in B2 compared to B1 ($\beta = -1.55$, $z = -5.93$, $p < 0.001$, Figure 11a).

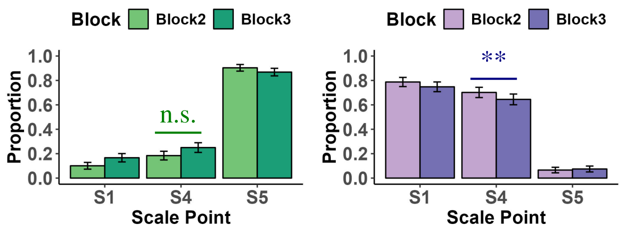


(a) Predication accepted. (b) Predication rejected.

Figure 11: Experiment 2: Block 1-2 comparison results.

To examine whether the observed effects persisted post-dialogue, B2 and B3 were compared using the same coding and modeling approach as in Experiment 1. Model outputs showed no significant change in acceptance rates at S4 ($\beta = 0.63$, $z = 1.41$, $p > 0.1$, Figure 12a), while rejection rates significantly decreased ($\beta = -0.62$, $z = -2.65$, $p < 0.01$, Figure 12b). Finally, following the same methodology used in Experiment 1, we compared B1 and B3. This analysis revealed a significant drop in acceptance rates ($\beta = 0.22$, $z = -3.58$, $p < 0.001$) and no difference in rejection rates ($\beta = 0.28$, $z = 1.47$, $p > 0.1$) in B3 compared to B1.

[2]The same analysis procedure was applied to the controls, yielding parallel patterns as in Experiment 1.



(a) Predication accepted. (b) Predication rejected.

Figure 12: Experiment 2: Block 2-3 comparison results.

Discussion

The comparison between B1 and B2 shows that participants' exposure to the precise bot drove participants to raise their SoPs in B2, as evidenced by higher rejection rates of the predication in B2 (i.e., *Agree*-responses) compared to B1 (i.e., *B*-responses), along with significantly lower acceptance rates in B2 (i.e., *Disagree*-responses) than in B1 (i.e., *A*-responses). While Experiment 2 results can be accommodated by Hypothesis 1, this hypothesis cannot account for the weakening effects observed in Experiment 1. Overall, the combined findings from both experiments are better accounted for by Hypothesis 2, which posits bidirectional SoP modulation.

The comparison between B2 and B3 revealed mixed patterns regarding whether the observed adaptation effects were confined to the dialogue. Specifically, acceptance rates for imprecise predications (i.e., *Disagree*-responses in B2 and *A*-responses in B3) were comparable, suggesting that the raised SoPs persisted. However, rejection rates (i.e., *Agree*-responses in B2 and *B*-responses in B3) decreased significantly, indicating a partial weakening reversion after the dialogue. This pattern is further supported by the B3–B1 comparison: acceptance rates (i.e., *A*-responses) significantly declined in B3 relative to baseline B1, while no significant difference was found in rejection rates (i.e., *B*-responses). Theses findings suggest that while raised SoPs endure to some extent, such upward adaptation partially reverts after the dialogue.

General Discussion

Existing work on imprecision has primarily focused on the static constraints that modulate its acceptability—such as contextual stakes, agent goals and speaker identity (Aparicio Terrasa, 2017; Beltrama & Schwarz, 2024; Burnett, 2014; Leffel et al., 2016; Mathis & Papafragou, 2022; Ronderos et al., 2024; Solt et al., 2017; Syrett et al., 2010; a.o.)—rather than the dynamic discourse factors that regulate its tolerance (though see Wu & Aparicio, 2025). The present study constitutes a first step towards filling this gap. In two studies, we investigate whether comprehenders adapt to the (im)precision preferences displayed by their interlocutor. Our experimental paradigm consisted of first-person dialogue experiments, where participants engage in real-time, implicit negotiations about the SoP with another interlocutor.

Experiment 1 results show that exposure to an imprecise speaker resulted in increased tolerance for imprecision. In contrast, Experiment 2 demonstrated that exposure to a precise speaker caused participants to become less tolerant of imprecision compared to baseline. The SoP strengthening effects detected in Experiment 2 provide partial support to our first hypothesis (H1), which states that adaptation effects should only be detected in the form of precisifications (Klecha, 2018; Lewis, 1979). However, H1 cannot accommodate Experiment 1 results, which clearly show that participants lowered their SoPs after exposure to an imprecise speaker. Taken together, Experiments 1 and 2 demonstrate that comprehenders dynamically adjust their standards of precision (SoP) based on their interlocutor's preferences. This pattern of results is more consistent with the second hypothesis (H2), which posits that SoP adaptation is bidirectional—such that exposure to both higher and lower SoPs leads to corresponding updates. More generally, our findings suggest that adaptation functions as a core mechanism of pragmatic reasoning, allowing comprehenders to dynamically update meaning representations in order to coordinate on utterance interpretation.

Furthermore, our results show that SoP adaptation effects persisted beyond the dialogue, suggesting that comprehenders do not simply align with their partner's preferences during linguistic interactions but, to some extent, internalize these adjustments beyond the immediate discourse. The weakening of the SoP remained stable in the post-dialogue phase, as evidenced by comparable acceptance and rejection rates between B2 and B3. However, in Experiment 2, raised SoPs partially reverted, with rejection rates of imprecise utterances decreasing post-dialogue. This suggests that tolerance for imprecision may be more stable than increased precision, raising new questions about the durability of such adaptations in dynamic meaning negotiation.

Finally, the current results highlight the importance of first-person engagement in studying linguistic coordination during conversation. Our findings suggest that first-person dialogue paradigms are an effective tool for probing dynamic meaning negotiations, offering a unique perspective on how speakers coordinate in real time through adaptive behavior.

Conclusion

This paper has explored the discourse dynamics of imprecision, focusing on how comprehenders adapt their SoPs based on their interlocutor's (im)precision preferences. Our results show that participants flexibly adapted towards the SoP displayed by their interlocutor: exposure to an imprecise speaker increased participants' tolerance for imprecision, while exposure to a precise interlocutor caused participants to adopt a stricter threshold. Furthermore, these effects did not immediately reverse once the conversation concluded. Methodologically, the current results underscore the effectiveness of first-person paradigms in the study of discourse dynamics phenomena.

References

Aparicio, H., Xiang, M., & Kennedy, C. (2015). Processing gradable adjectives in context: A visual world study. In *Semantics and Linguistic Theory (SALT)* (Vol. 25, pp. 413–432).

Aparicio Terrasa, H. (2017). *Processing context-sensitive expressions: The case of gradable adjectives and numerals* (Doctoral dissertation). The University of Chicago.

Barker, C. (2002). The dynamics of vagueness. *Linguistics and Philosophy*, *25*(1), 1–36.

Barker, C. (2013). Negotiating taste. *Inquiry*, *56*(2-3), 240–257.

Beltrama, A., & Schwarz, F. (2021). Imprecision, personae, and pragmatic reasoning. In *Semantics and Linguistic Theory (SALT)* (Vol. 31, pp. 122–144).

Beltrama, A., & Schwarz, F. (2022). Social identity, precision and charity: when less precise speakers are held to stricter standard. In *Semantics and Linguistic Theory (SALT)* (Vol. 32, pp. 575–598).

Beltrama, A., & Schwarz, F. (2024). Social identity affects imprecision resolution across different tasks. *Semantics and Pragmatics*, *17*, 10–EA.

Beltrama, A., Solt, S., & Burnett, H. (2023). Context, precision, and social perception: A sociopragmatic study. *Language in Society*, *52*(5), 805–835.

Burnett, H. (2014). A delineation solution to the puzzles of absolute adjectives. *Linguistics and Philosophy*, *37*, 1–39.

Heim, S., Peiseler, N., & Bekemeier, N. (2020). "few"' or "many"'? an adaptation level theory account for flexibility in quantifier processing. *Frontiers in psychology*, *11*, 382.

Horn, L. R. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Kao, J. T., Wu, J. Y., Bergen, L., & Goodman, N. D. (2014). Nonliteral understanding of number words. In *Proceedings of the national academy of sciences* (Vol. 111, p. 12002-12007).

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, *30*, 1–45.

Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, *81*(2), 345–381.

Klecha, P. (2018). On unidirectionality in precisification. *Linguistics and Philosophy*, *41*, 87–124.

Krifka, M. (2002). Be brief and vague! and how bidirectional optimality theory allows for verbosity and precision. In D. Restle & D. Zaefferer (Eds.), *Sounds and systems: Studies in structure and change. A Festschrift for Theo Vennemann* (pp. 439–458). Berlin, New York: De Gruyter Mouton.

Krifka, M. (2007). *Approximate interpretation of number words*. Berlin: Humboldt-Universität zu Berlin, Philosophische Fakultät II.

Lasersohn, P. (1999). Pragmatic halos. *Language*, *75*(3), 522–551.

Lauer, S. (2012). On the pragmatics of pragmatic slack. In *Proceedings of Sinn und Bedeutung* (Vol. 16, pp. 389–402).

Lauer, S. (2013). *Towards a dynamic pragmatics* (Doctoral dissertation). Stanford University.

Leffel, T., Xiang, M., & Kennedy, C. (2016). Imprecision is pragmatic: Evidence from referential processing. In *Semantics and Linguistic Theory (SALT)* (Vol. 26, pp. 836–854).

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, *8*, 339–359.

Mathis, A., & Papafragou, A. (2022). Agents' goals affect construal of event endpoints. *Journal of Memory and Language*, *127*, 104373.

Pecsok, E., & Aparicio, H. (2024). How can they both be right?: Faultless disagreement and semantic adaptation. In *Proceedings of the annual meeting of the cognitive science society* (pp. 3939–3945).

Ronderos, C. R., Noveck, I., & Falkum, I. L. (2024). Straight enough: Deriving imprecise interpretations of maximum standard absolute adjectives. *Glossa Psycholinguistics*, *3*(1), 1-36.

Rotstein, C., & Winter, Y. (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics*, *12*, 259–288.

Schuster, S., & Degen, J. (2019). Speaker-specific adaptation to variable use of uncertainty expressions. In *Proceedings of the annual meeting of the cognitive science society* (pp. 2769–2775).

Schuster, S., & Degen, J. (2020). I know what you're probably going to say: Listener adaptation to variable use of uncertainty expressions. *Cognition*, *203*, 104285.

Solt, S. (2015). Vagueness and imprecision: Empirical foundations. *Annu. Rev. Linguist*, *1*(1), 107–127.

Solt, S., Cummins, C., & Palmović, M. (2017). The preference for approximation. *International Review of Pragmatics*, *9*(2), 248–268.

Syrett, K., Kennedy, C., & Lidz, J. (2010). Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics*, *27*(1), 1–35.

van Der Henst, J., Carles, L., & Sperber, D. (2002). Truthfulness and relevance in telling the time. *Mind & Language*, *17*(5), 457–466.

Wu, Y., & Aparicio, H. (2025). Disagreements do not automatically raise the standard of precision. In *Proceedings of Experiments in Linguistic Meaning* (Vol. 3, p. 435-446).

Xiang, M., Kramer, A., & Kennedy, C. (2020). Semantic adaptation in gradable adjective interpretation [Unpublished Manuscript].

Yildirim, I., Degen, J., Tanenhaus, M. K., & Jaeger, T. F. (2016). Talker-specificity and adaptation in quantifier interpretation. *Journal of memory and language*, *87*, 128–143.

Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments (IBEX)*.