The effect of intonation on scalar and ignorance inferences

Brian Buccola and Daniel Goodhue*
Michigan State University and Zentrum Allgemeine Sprachwissenschaft

1 Overview

We present the results of two experiments investigating the effect of intonation on the interpretation of indirect answers to polar questions. We compare how the English rise-fall-rise contour and neutral falling contour affect scalar inferences and ignorance inferences. Our results suggest that (i) both contours are compatible with scalar inferences, but that (ii) only rise-fall-rise is also compatible with ignorance inferences. We propose an account of the data that combines a classic uncertainty view of rise-fall-rise meaning (Ward & Hirschberg 1985) with a formal pragmatics that requires both scalar and ignorance inferences to be expressed in grammar (Meyer 2013; Fox 2016; Buccola & Haida 2019). We show that our account can also explain results from another recent experiment that seem to conflict with our own, and we point out empirical challenges for other proposals for the meaning of rise-fall-rise.

2 Introduction and background

Ward & Hirschberg (1985) analyze the English rise-fall-rise contour (L*+H L-H%) (RFR) as a marker of uncertainty. (1) and (2) demonstrate two different ways that RFR can express uncertainty in indirect answers to polar questions (rather than a direct "yes" or "no" answer).

(1) A: Have you ever been west of the Mississippi?
B: I've been to Missouri. (Ward & Hirschberg 1985:767)
L*+H L-H%

(2) A: Did she get a hundred on the midterm?
B: She got a ninety-eight. (Ward & Hirschberg 1985:767)
L*+H L-H%

In (1), B can use RFR to convey uncertainty about whether Missouri is west of the Mississippi, and thus uncertainty about whether the answer to A's question is "yes" or "no". In (2), B conveys that the person in question did not get a hundred on the midterm, and thus that the answer to A's question is "no", but B nevertheless conveys uncertainty about something else. Ward & Hirschberg write, "B might

*Thanks to Emmanuel Chemla, Alex Göbel, Eszter Ronai, audiences at the Chicago Linguistic Society 59, members of the MSU Semantics and Pragmatics Lab, Manfred Krifka's Speech Acts in Grammar and Discourse research group at ZAS, and members of Jennifer Cole's lab at Northwestern University. Daniel Goodhue is supported by the ERC Advanced Grant 787929 "Speech Acts in Grammar and Discourse" (SPAGAD).

be uncertain about whether ninety-eight and a hundred are close enough (on the cardinal scale) for A's purposes" (p. 767).

(1) and (2) can be related to examples of ignorance inference (II) and scalar inference (SI), as in (3).[1] If B uses RFR (left version), then two interpretations are intuitively possible: (i) B conveys uncertainty about whether all of the guests ate dinner, hence about whether the answer to A's question is "yes" or "no". In other words, RFR may lead the listener to draw an II.[2] (ii) B conveys that not all of the guests ate dinner, hence that the answer to A's question is "no", while also conveying uncertainty about something else, such as whether the answer that some but not all of the guests ate dinner is a satisfactory answer for A's purposes. In other words, RFR is compatible with the listener drawing a SI. (For an overview of SIs and IIs, see Geurts 2010, among many others.)

(3) A: Did all of the guests eat dinner?

B: *Some* of them ate dinner. — $\underbrace{\text{L*+H} \qquad\qquad \text{L-H\%}}_{\text{RFR}}$

B: *Some* of them ate dinner. — $\underbrace{\text{H*} \qquad\qquad \text{L-L\%}}_{\text{Fall}}$

By contrast, if B uses neutral falling intonation (H* L-L%) (Fall) (right version), only a SI interpretation seems to be available. If B intends to convey an II, the use of Fall in (3) is intuitively odd. In the general discussion, we will explore two different explanations for this restriction on Fall, and argue in favor of a view in which Fall is a meaningless default intonation, and the restriction is explained via the formal pragmatics of IIs.

If the intuitions just reported hold, then if a speaker intends to convey an II in a context like (3), a listener should reasonably expect them to use RFR rather than Fall. On the flip side, since RFR is compatible with both SIs and IIs, if a listener hears RFR in a context like (3), they cannot know which inference the speaker intends to convey. Still, the chances that they intend to convey an II are higher with RFR than Fall since we have claimed the latter cannot do so. Moreover, we have the intuition that, if confronted with a dialogue like (3) with RFR in B's reply, an II is more likely to be drawn than a SI. Thus, our initial hypothesis is as in (4):

(4) *Hypothesis 1.* In indirect answers with weak scalar items replying to polar questions with strong scalar items, RFR will **increase** the rate of **IIs** that listeners draw relative to Fall.

It will be useful to distinguish our hypothesis from two other logically possible ones. The three jointly exhaust the logical hypothesis space.[3]

[1] Sound files for (3) are available at `https://doi.org/10.17605/OSF.IO/76TZF`.

[2] While we use "uncertainty" to characterize the meaning of RFR and "ignorance" to identify a kind of inference about strong scalemates that is distinct from a scalar inference, we take the two terms to refer to a single epistemic state in which an agent neither believes a proposition is true ($\neg\mathcal{B}p$) nor believes that it is false ($\neg\mathcal{B}\neg p$). That is, uncertainty = ignorance = $\neg\mathcal{B}p \wedge \neg\mathcal{B}\neg p$.

[3] One crucial assumption that constrains this hypothesis space is that we assume that indirect answers containing *some* to polar questions containing *all* necessarily result in either an II or a SI about the stronger *all*-alternative. That is, the answer either results in an "I don't know" answer to the question, or a "no" answer, but never a "yes" answer. Thus, if RFR increases IIs relative to Fall, it

(5) *Hypothesis 2*. In indirect answers with weak scalar items replying to polar questions with strong scalar items, RFR will **increase** the rate of **SIs** that listeners draw relative to Fall.

(6) *Hypothesis 3 (null hypothesis)*. In indirect answers with weak scalar items replying to polar questions with strong scalar items, RFR will have **no effect** on the rate of SIs and IIs that listeners draw relative to Fall (that is, there will be no difference between RFR and Fall).

Hypothesis 3, which is also the null hypothesis in that it predicts that intonation will have no effect on the rate of IIs/SIs in these contexts, seems possible given that we have already seen that RFR is in principle compatible with both SIs and IIs. We may fail to reject the null if, for example, there is a strong bias to draw SIs regardless of intonation.

Hypothesis 2, on the other hand, seems unexpected: given that RFR conveys uncertainty, and that we assume that Fall is not able to convey IIs in contexts like (3), we do not expect RFR to lead listeners to draw more SIs than Fall.

However, de Marneffe & Tonhauser (2019) report a perception experiment that they claim shows that RFR *increases* the rate of SIs relative to Fall, in support of the counterintuitive Hypothesis 2. We are skeptical of this conclusion, since the design of the experiment does not distinguish SIs from IIs: participants were only asked whether or not B's utterance conveys a "yes" answer to A's polar question containing the stronger scalemate. Results showed that RFR *decreased* the rate at which B's utterance was judged to be a "yes" answer to A's question relative to Fall. Crucially, if a participant judges B's utterance not to count as a "yes" answer to A's question, that is consistent with the participant having drawn *either a SI or an II*, since either one contextually implies that it is not the case that B is giving a "yes" answer to A's question. (See Westera *et al.* 2020 for a similar critique of de Marneffe & Tonhauser 2019.)

In recent and ongoing work, Göbel & Ronai (2023) report the results of an experiment that is similar to de Marneffe & Tonhauser's, but that successfully distinguishes SIs from IIs: participants were asked whether or not B's utterance conveys a "no" answer to A's polar question containing a stronger scalemate. If participants gave a positive answer to this test question, it means they drew a SI and took B's utterance to convey a "no" answer. If they gave a negative answer, it is compatible with either an "I don't know" II or a "yes" answer to A's question. Göbel & Ronai find that RFR increases the rate of participants' positive answers to the experimental question, and hence that it increases the rate of SIs relative to Fall, in support of Hypothesis 2. These results are puzzling from the perspective of an uncertainty

implies that it decreases SIs, and vice versa. This is intuitively plausible in our examples containing logical scalar pairs like ⟨*some*, *all*⟩; however, it is less plausible in question-answer pairs containing gradable adjective pairs of the sort studied by both de Marneffe & Tonhauser (2019) and Göbel & Ronai (2023), e.g. ⟨*strenuous*, *exhausting*⟩ and ⟨*happy*, *ecstatic*⟩. In the latter cases, a reply with a weak scalemate to a polar question with a strong scalemate could be taken to indicate an indirect "yes" answer, as well as "no" and "I don't know"; thus, RFR could possibly increase both SIs and IIs relative to Fall. We return to these points in the general discussion. Moreover, another limiting factor in our study is that we only compare RFR and Fall, and ignore other possible contours.

theory of RFR such as Ward & Hirschberg's. We will return to Göbel & Ronai's results in the general discussion.

In the following two sections, we report on two experiments designed to directly differentiate SIs and IIs.[4] Experiment 1 tests our intuition about Hypothesis 1, and also provides a chance to see if the results in Göbel & Ronai 2023 that support Hypothesis 2 will emerge with a slightly different experimental paradigm. Experiment 2 tests a revised hypothesis.

3 Experiment 1

3.1 Design and methods

Participants listened to dialogues and answered a multiple choice question about each. They were told to treat each dialogue as independent from all others. In each target trial, participants were presented with a dialogue like the one in (3). To produce the stimuli for these dialogues, the two authors (both male, native speakers of American English) were recorded in a sound-attenuated booth. The experimental manipulation was of the intonation of B's reply to A's question: B's utterance was produced either with RFR or with Fall. Thus, the experiment had one independent variable, INTONATION, with two conditions, RFR and FALL. The recordings were examined using Praat (Boersma & Weenink 2023) to ensure that they featured the intended contours. The dialogues were not presented visually in order to avoid any potential influence of visual punctuation. The dialogue played automatically at the start of the trial. Participants, who participated virtually, were requested to use headphones and could replay the dialogue as many times as they liked.

Participants then saw a question on the screen with two possible answers:

(7) Based on Mason's response, which of the following sentences is true?
  a. Mason thinks that not all of the guests ate dinner.
  b. Mason isn't sure whether or not all of the guests ate dinner.

These answers were displayed below the question and arranged vertically, as in (7), order pseudo-randomized. There were 8 items, and the experiment was run within participant, so there were 16 test observations per participant.

There were also 16 filler trials per participant that were used to distract from the purpose of the task and to ensure participants had paid attention. 8 of the fillers had unambiguously correct answers; if a participant gave any incorrect answers, they were removed. An example of such a filler can be found in (8), where the answer "No" was produced with one of two intonation contours, Fall or Fall-Rise:[5]

(8) A: Are you a friend of Jenny's?
  B: No.

[4] All materials for both experiments, including all stimuli recordings, are freely available at `https://doi.org/10.17605/OSF.IO/76TZF`. While we make these stimuli freely available for examination by other researchers, we request that others not use them in an experiment without contacting us for consent.

[5] The filler sound files and materials come from an experiment reported by Goodhue & Wagner (2018), who claim that this particular fall-rise is the contradiction contour (Liberman & Sag 1974) and is distinct from RFR.

Like in (7), the participant was asked which of two sentences was true (e.g., whether B was a friend of Jenny's, or not). Test and filler trials were pseudo-randomized. The experiment began with 2 practice trials, one like a test trial, and the other like a filler, to familiarize participants and allow them to adjust their audio. Thus, each participant provided 34 observations. The experiment took less than 10 minutes.

3.2 Participants

35 participants, self-identified as native speakers of North American English and located in the US, were recruited via Mechanical Turk. 3 participants were eliminated due to providing an incorrect answer to 1 out of 8 unambiguous fillers. 2 more participants were excluded for responding before they had completed listening to the audio recordings in every trial. Thus, we report results for 30 participants.

3.3 Predictions

Recall Hypothesis 1 from Section 2:

(9) *Hypothesis 1*. In indirect answers with weak scalar items replying to polar questions with strong scalar items, RFR will **increase** the rate of **IIs** that listeners draw relative to Fall.

Hypothesis 1 could fail to be confirmed by finding results similar to Göbel & Ronai 2023, in which RFR increases the rate of SIs relative to Fall (cf. Hypothesis 2). Both Hypotheses 1 and 2 could fail to be confirmed if intonation has no effect on inferences (RFR and Fall produce the same rates of IIs and SIs) (cf. Hypothesis 3).

3.4 Results

See Figure 1. Initial inspection of the plot on the left shows that participants derived SIs for 92% of Fall utterances and for 80% of RFR utterances. Thus, participants were overwhelmingly more likely to draw SIs than IIs overall. At the same time, there was a numerical difference in the direction of Hypothesis 1 (more IIs for RFR than Fall). However, closer inspection of the righthand plot reveals that the difference is driven by 5 outliers: a plurality of participants chose SIs in all trials (the 8/8 row), while 5 participants were more likely to draw IIs than SIs in the RFR condition (bottom right of rightmost panel, rows 0/8 to 3/8).

A mixed effects logistic regression with Intonation as a fixed effect and random intercepts and slopes for participant and item revealed no effect of intonation ($\hat{\beta}$ = .010, *se* = 1.152, $p > .99$). Thus, we cannot reject the null hypothesis (Hypothesis 3) in favor of either Hypothesis 1 or Hypothesis 2.

3.5 Discussion

When SI and II are two distinct options, RFR does not increase the rate of SI, nor does it increase the II rate. Instead SI is almost always chosen: 21 participants chose SI in over 87% of trials, with 14 of them choosing SI in all trials. It seems that there is a strong bias for drawing SIs in our experimental contexts, which resulted in a ceiling effect in which participants chose SI almost all of the time. Given this ceiling effect, while the result fails to reject the null hypothesis, Hypothesis 3, it

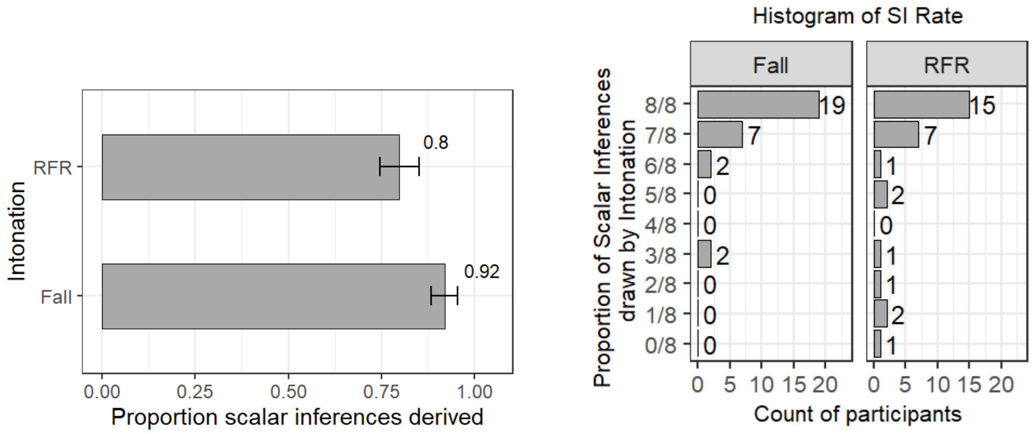

Figure 1: Experiment 1 data visualizations. Left plot with 95% CIs.

is possible that either Hypothesis 1 or Hypothesis 2 is on roughly the right track. Viewed from the perspective of Hypothesis 1, it seems that many participants either ignored intonation entirely, or took RFR to convey uncertainty about something other than the stronger alternative, like whether the speaker's response satisfies a more general question under discussion. But a different test might still reveal that RFR is more appropriate for conveying IIs than Fall. Viewed from the perspective of Hypothesis 2, participants may have had independent reasons to draw SIs in virtually all cases. But a different test might still reveal that RFR is in a sense more appropriate for conveying SIs than Fall.

To further probe the phenomenon, in Experiment 2 we forced participants to associate one contour with a SI and another with an II. We effectively stipulated that B intends to convey an II in one dialogue and a SI in another, and we asked participants which intonation is more appropriate for which inference. If RFR conveys uncertainty better than Fall, then it should be associated with IIs at a higher rate. In other words, while RFR may not increase II rates relative to Fall (given that both capably convey SIs), RFR may nevertheless be more appropriate than Fall at conveying IIs when the speaker has the intention to convey an II.

4 Experiment 2

4.1 Design and methods

Experiment 2 used the stimuli from Experiment 1, but paired both levels of an item in a single trial. So in one trial, participants heard two dialogues in a row labeled "Version 1" and "Version 2", order pseudo-randomized. These dialogues differed only in contour, RFR vs. FALL, they were not presented visually, and unlimited replays were available. Participants then saw the following question with two possible answers underneath, arranged vertically, order pseudo-randomized:

(10) Which of the following two options best describes the meanings of the two responses?

a. In version 1, Mason thinks that not all of the guests ate dinner, and in version 2, Mason isn't sure whether or not all of the guests ate dinner.

b. In version 1, Mason isn't sure whether or not all of the guests ate dinner, and in version 2, Mason thinks that not all of the guests ate dinner.

In essence, then, participants were asked to match one of the contours with a SI and the other with an II. While intonation is the only factor differing across the two dialogues, the experiment technically does not have an independent variable, since both contours are present in every trial. Instead, we are looking for whether participant responses are different from chance, and if so, in which direction.

Since there were two dialogues per trial, there were 8 test observations per participant in Experiment 2. The fillers in Experiment 2 were identical to Experiment 1, except only the unambiguous fillers were included, thus 16 trials total per participant, pseudo-randomized, plus 2 practice trials.

4.2 Participants

36 adults with the same attributes from Experiment 1 were recruited through Mechanical Turk. This time, 3 were excluded due to missing fillers.

4.3 Predictions

The experiment tests Hypothesis 4, building on Hypothesis 1, based on the uncertainty view of RFR:

(11) *Hypothesis 4*. In indirect answers with weak scalar items replying to polar questions with strong scalar items, when forced to associate one of RFR and Fall with a SI and the other with an II, RFR will be associated with **IIs** at a higher rate than Fall by listeners.

On the other hand, if the results are more in keeping with the de Marneffe & Tonhauser 2019 and Göbel & Ronai 2023 view, then a hypothesis opposite from Hypothesis 4 will be confirmed, namely one in which RFR will be associated with **SIs** at a higher rate than Fall (Hypothesis 5). Finally, another possibility is that intonation may again produce no effect, the null hypothesis, which in this case would mean that participants will choose an association at chance, 50% (Hypothesis 6).

4.4 Results and discussion

See Figure 2. The plot on the left shows that participants associated falling intonation with SI and RFR with II in 68% of trials (and therefore made the reverse association in 32% of trials).

An intercept-only mixed effects logistic regression with random intercepts for participant and item revealed that the result was significantly different from if the answers had been chosen at chance ($\hat{\beta}$ = .932, se = .278, $p < .001$). In particular, participants were much more likely to associate RFR with II and Fall with SI than vice versa, allowing for the rejection of the null Hypothesis 6 in favor of Hypothesis 4, which was motivated by the uncertainty view of RFR.

The righthand plot further suggests that the participants primarily fell into two groups. The Y axis contains 9 bins that count in how many trials participants associated RFR with II and Fall with SI. For the purposes of this discussion, we combine

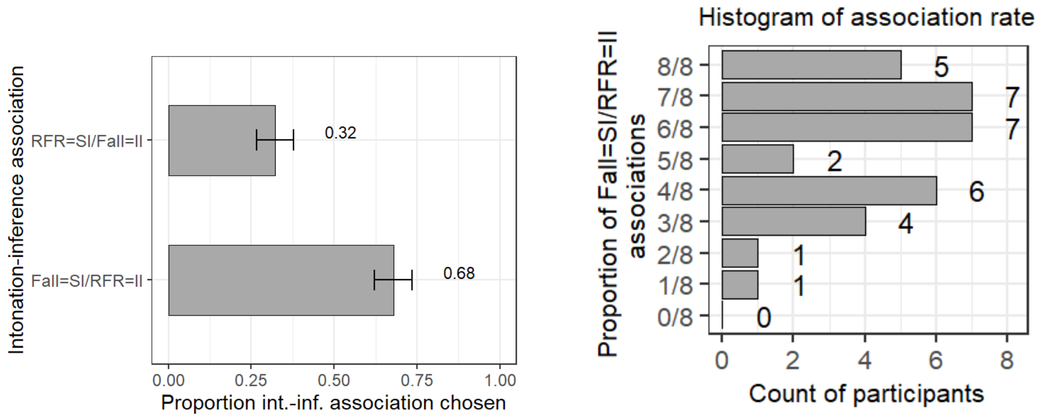

Figure 2: Experiment 2 data visualization. Left plot with 95% CIs.

these bins into three groups: the first group (6/8, 7/8, and 8/8) contains a majority of participants (n=19), who were more likely to associate RFR with II and Fall with SI. This group drove the result. Another group (3/8, 4/8, and 5/8; n=12) clustered around chance. For these, the experimental manipulation seems not to have had an effect; that is, they seemed to ignore intonation. The final group (0/8, 1/8, and 2/8) only contains two participants, suggesting that very few speakers reflect the view that RFR increases the rate of SIs relative to Fall (Hypothesis 5).

5 General discussion

Our results from Experiment 1 show that RFR is consistent with SIs, and in fact that participants overwhelmingly draw SIs, and not IIs, regardless of intonation. We wanted to see if one of the intonations was preferred for expressing IIs, so in Experiment 2 we gave participants both intonations at the same time and asked them to associate one with an II and the other with a SI. We found that most participants were significantly more likely to associate RFR with II and Fall with SI, with another group of participants who were not affected by the manipulation.

We will explore two approaches to explaining these results. The first has three components: (i) a strong uncertainty semantics for RFR, (ii) a non-uncertainty (possibly vacuous) semantics for Fall, and (iii) a formal pragmatics that derives ignorance inferences as grammatical entailments (Meyer 2013; Fox 2016; Buccola & Haida 2019). The second approach combines a strong certainty (or exhaustification) semantics for Fall with various non-uncertainty views of RFR. We argue in favor of the first approach. We then discuss how to reconcile the seeming contradiction between our experimental results and those of Göbel & Ronai (2023).

5.1 Conveying ignorance grammatically

We adopt Ward & Hirschberg's (1985) view that RFR conveys *uncertainty*, which we take to be equivalent to *ignorance* in the SI literature. Let $\mathcal{B}$ symbolize the attitude of belief. An agent is uncertain about a proposition p if and only if they neither believe p is true ($\neg\mathcal{B}p$) nor believe p is false ($\neg\mathcal{B}\neg p$). Thus, uncertainty can be symbolized as $\neg\mathcal{B}p \wedge \neg\mathcal{B}\neg p$. RFR conveys uncertainty about different propositions p in different contexts. We identified two in (3): a stronger *all*-alternative to a *some*-utterance, and a proposition p about whether the current utterance satisfies a larger goal, so that RFR conveys "I'm uncertain whether [$_p$ this utterance suffices]".

Our formal pragmatics begins with Fox's (2016) intuitive claim that silence about a relevant proposition p is uncooperative, including silence about one's ignorance about p. Fox takes this observation as evidence that relevance is closed under speaker belief (as well as negation and conjunction): if p is relevant, then so is $\mathcal{B}p$, and thus so is $\neg\mathcal{B}p \wedge \neg\mathcal{B}\neg p$, the speaker's ignorance about p. Since the Maxim of Quantity compels a speaker to settle (entail, or entail the negation of) any relevant proposition, it follows that if a speaker is ignorant about p, then the speaker is obliged to choose an utterance that entails their ignorance.[6] In a dialogue like (3), given A's question, the stronger *all*-alternative to B's *some*-utterance is relevant. Thus, if B is ignorant about *all*, then B's utterance must entail this ignorance. But Meyer (2013) argues that (even with the help of covert operators) there is no parse of a *some*-utterance that entails ignorance about *all*. Buccola & Haida (2019) share that view but claim that it holds only for *some*-utterances produced with Fall; they propose that RFR, on the other hand, can entail ignorance (assuming RFR targets the stronger alternative).[7] We follow Buccola & Haida and assume, crucially, that Fall is either a meaningless default intonation (as argued in Goodhue forthcoming), or at least, Fall's meaning does not entail uncertainty/ignorance (as seems likely).

We turn now to applying our views of RFR and ignorance to our experimental results. Experiment 1 found no effect of intonation, with a strong bias toward SIs. Our account does not predict SI bias in the RFR condition, but is consistent with it in that RFR does not *need* to target the stronger alternative *that all of the guests ate dinner*, but can instead target whether the current utterance satisfies a larger goal. This seems to be how Experiment 1's participants interpreted RFR. Our account predicts the SI bias in the Fall condition, since we claim that Fall cannot convey an II. In Experiment 2, most participants associated RFR with II and Fall with SI. Our account requires IIs to be conveyed grammatically, i.e. to be entailed. RFR is able to do so, while Fall is not. Thus, our account explains the majority behavior found in Experiment 2. As for the remaining participants, most chose at chance, i.e. intonation had no effect. One possible explanation is that the artificial experimental setting made some participants insensitive to the prosodic manipulation. Finally, only two participants chose the other association, Fall with II and RFR with SI. This follows from our view that Fall cannot grammatically express IIs.

For our account to work, the uncertainty expressed by RFR must be strong ($\neg\mathcal{B}p \wedge \neg\mathcal{B}\neg p$). It cannot be a mere lack of certainty ($\neg\mathcal{B}p$), which is consistent with (is entailed by) both uncertainty, and certain falsity ($\mathcal{B}\neg p$). Constant (2012) proposes such a weaker account in which RFR conveys that the speaker "cannot safely claim" an alternative p. One reason for not being able to claim p is that you are ignorant about it ($\neg\mathcal{B}p \wedge \neg\mathcal{B}\neg p$), but another is that you believe it to be false ($\mathcal{B}\neg p$). This meaning for RFR cannot be combined with our account, since ignorance must be grammatically conveyed. The issue is subtle but important: on

[6] See Buccola & Haida 2019 for discussion and a formal proof.

[7] Buccola & Haida (2019) identify other types of sentences that, under their theoretical assumptions, can entail ignorance: *At least some of the guests ate dinner*, *Some or all of the guests ate dinner*, and *Some of the guests ate dinner but I'm not certain whether they all did*.

the view that we have adopted, RFR entails uncertainty, but it is consistent with both IIs and SIs because that uncertainty can target different propositions. On Constant's view, RFR does not entail uncertainty, but instead a weaker lack of certainty that is directly consistent with both IIs and SIs; thus, it cannot do the work of encoding ignorance for the purpose of grammatically conveying IIs.

5.2 Exhaustifying intonationally

The mirror image of our explanation of the results would be to treat Fall as semantically encoding certainty or exhaustification, while RFR has some non-uncertainty semantics. Thus, Fall applied to an utterance of *Some of the guests ate dinner* in a context in which the stronger *all*-alternative is relevant would entail an exhaustive SI, *that not all of the guests ate dinner*. Like our account, such a view explains the SI bias in the Fall condition of Experiment 1, and is merely consistent with the SI bias in the RFR condition. It also explains the result of Experiment 2, since Fall directly conveys SIs, and so participants should not choose to associate it with IIs.

This view combines with any account of RFR: Constant (2012); Göbel (2019), who says RFR requires a salient higher alternative; Wagner (2012), who says RFR requires the performance of an alternative speech act to be justified; and Westera (2019), who says RFR conveys that a focus congruent QUD, but not the main QUD, is fully addressed. All of these accounts except ours require the Fall-as-exhaust assumption in order to explain our results. While a complete discussion awaits future work, the data point in (12), building on (2), speaks in favor of the uncertainty view, since RFR is odd in (12b) when there is no uncertainty to associate with it:

(12) A: Did she get a hundred on the midterm?
  a. B: She got a ninety-eight. (RFR)
  b. B: #She got a sixty. (RFR) (Ward & Hirschberg 1985)

Wagner (2012) and Westera (2019) may predict (12b), but Constant (2012) and Göbel (2019) incorrectly predict RFR to be acceptable in any case of SI, including (12b).

Here are two drawbacks to Fall-as-exhaust. (i) Fall is not necessary to achieve exhaustification: SIs arise from RFR utterances, and they may also arise from other utterances with rising intonations, like rising declaratives and the contradiction contour. (ii) Fall is not sufficient for producing exhaustification. A SI can be canceled or suspended even when the utterance that triggers it was produced with Fall, e.g. *Some of the guests brought wine. In fact, all of them did*, or *Some of the pears are bruised. I wonder if they all are*. Such examples pose a serious challenge for the view that Fall semantically expresses exhaustification. But if Fall-as-exhaust is incorrect, then the non-uncertainty theories of RFR that we cited above cannot explain the results of Experiment 2, as far as we can tell.

5.3 Reconciling conflicting results

Whereas our experiment probed the $\langle \textit{some}, \textit{all} \rangle$ scale only, de Marneffe & Tonhauser (2019) and Göbel & Ronai (2023) investigated a greater variety of scalar pairs, many adjectival (see fn. 3). This makes a third inference beyond II and SI available in their stimuli: the utterance of the weak alternative could be taken by participants as an indirect "yes" answer to the polar question containing the stronger alternative, as in (13c) (cf. discussion in Ronai & Xiang 2022). This interpretation may be more apparent with prosody that conveys positive emotional affect, e.g. a higher maximum pitch/larger pitch excursion (cf. the intuition that "intensification" may correlate with a "yes" answer in Ronai & Xiang 2023).

(13) A: Was the winner ecstatic?
B: She was happy.
a. $\rightsquigarrow$ She was not ecstatic. "No" - SI
b. $\rightsquigarrow$ Uncertain whether or not she was ecstatic. "I don't know" - II
c. $\rightsquigarrow$ She was ecstatic. "Yes" - Strong Alt. (SA)

The participants in Göbel & Ronai's study were recorded producing B's utterance in dialogues like (13). Then they were asked, "Given your response, do you think A would conclude that the winner was not ecstatic?" If they gave a positive answer, it meant they thought A would draw a SI and conclude a "No" answer like (13a) (and we assume this means that participants themselves intended a SI when they produced the utterance). If they gave a negative answer to the test question, it was consistent with both an II/"I don't know" answer like in (13b) and a SA/"Yes" answer to A's question like in (13c). They found that after participants had produced RFR, the SI rate was 70%, while after Fall it was only 45%.

Despite the apparent contrast between this result and our Experiment 2, we believe that our uncertainty account of RFR can explain this result. Before speaking, participants had to decide whether they thought they were giving a "Yes", "No", or "I don't know" answer to A's question, i.e. which of the three inferences in (13) they intended. RFR is appropriate for conveying both "No"/SI (13a) and "I don't know"/II (13b), as discussed for (3). But RFR is intuitively inappropriate as a "Yes"/SA answer (13c) since there is no obvious uncertainty being conveyed. By contrast, Fall is intuitively appropriate for "Yes"/SA (13c) and "No"/SI (13a) (again, the "Yes"-Fall may be more natural with a higher H* pitch accent than the "No"-Fall), but not "I don't know"/II (13b). Thus, Fall was likely the primary means for conveying SA answers. Assuming that participants frequently intended such answers, only 45% of Falls were SIs because the rest were SAs. Assuming also that there was a bias for SIs and against IIs like in our Experiment 1, we understand why the SI rate was higher for RFR: RFR could not convey SAs, only SIs and IIs, so the SI bias means that most RFRs were SIs.[8]

The preceding discussion is one possible explanation for why RFR had a much higher SI rate than Fall in this experiment. One data point of interest that has not been reported yet from this experiment is, focusing just on the SIs (the positive answers to the test question), what proportion of the productions were RFR vs. Fall? In other words, when naïve speakers had the intention to convey a SI, which contour were they more likely to reach for and by how much? This latter question highlights

[8]Given the way the test question was posed, it is in fact possible that participants intended IIs with their RFR utterances more than the SI rate reflects. That is, after producing a RFR-II utterance, some participants may have decided upon reflection that the addressee might nevertheless draw a SI, and answered the test question accordingly, pushing up the SI rate.

the need for a future study that asks participants to produce indirect answers with a specific inference intended (SI vs. II (vs. SA)).

6 Conclusion

We designed two experiments that directly tease apart SIs from IIs. We found that (i) participants happily draw SIs regardless of intonation, but that (ii) only RFR is compatible with IIs. We take our results to be a vindication of the idea that RFR conveys uncertainty, despite the general preference to draw SIs over IIs. In future research, it would be fruitful to isolate different intended inferences, so as to observe which intonation participants prefer given that intended inference. We also hope to look at different theories of RFR more systematically, comparing their predictions against both our and others' data in the rich RFR literature.

References

Boersma, P., & D. Weenink, 2023. Praat: Doing phonetics by computer [computer program]. Version 6.3.10.

Buccola, B., & A. Haida. 2019. Obligatory irrelevance and the computation of ignorance inferences. *Journal of Semantics* 36. 583–616.

Constant, N. 2012. English rise-fall-rise: A study in the semantics and pragmatics of intonation. *Linguistics and Philosophy* 35. 407–442.

de Marneffe, M.-C., & J. Tonhauser. 2019. Inferring meaning from indirect answers to polar questions. In *Questions in Discourse: The Contribution of the Rise-Fall-Rise Contour*, 132–163. Brill.

Fox, D., 2016. On why ignorance might be part of literal meaning. Commentary on Marie-Christine Meyer, MIT Workshop on Exhaustivity.

Geurts, B. 2010. *Quantity Implicatures*. Cambridge University Press.

Göbel, A. 2019. Additives pitching in: L*+H signals ordered focus alternatives. In *Proceedings of Semantics and Linguistic Theory (SALT)*, volume 29, 279–299.

Göbel, A., & E. Ronai, 2023. On the meaning of intonational contours: A view from scalar inference. Abstract for Semantics and Linguistic Theory (SALT) 33.

Goodhue, D. forthcoming. Everything that rises must converge: Toward a unified account of inquisitive and assertive rising declaratives. In *Biased Questions: Experimental Results and Theoretical Modelling*, 1–30. Language Science Press.

Goodhue, D., & M. Wagner. 2018. Intonation, *yes* and *no*. *Glossa: A Journal of General Linguistics* 3. 1–45.

Liberman, M., & I. Sag. 1974. Prosodic form and discourse function. In *Proceedings of the Chicago Linguistic Society (CLS)*, volume 10, 416–427.

Meyer, M.-C. 2013. *Ignorance and Grammar*. Massachusetts Institute of Technology dissertation.

Ronai, E., & M. Xiang. 2022. Quantifying semantic and pragmatic effects on scalar diversity. In *Proceedings of the Linguistic Society of America (LSA)*, volume 7, p. 5216.

Ronai, E., & M. Xiang, 2023. What could have been said? Alternatives and variability in pragmatic inferences. Manuscript, Northwestern University and University of Chicago.

Wagner, M. 2012. Contrastive topics decomposed. *Semantics and Pragmatics* 5. 1–54.

Ward, G., & J. Hirschberg. 1985. Implicating uncertainty: The pragmatics of fall-rise intonation. *Language* 61. 747–776.

Westera, M. 2019. Rise-fall-rise as a marker of secondary QUDs. In *Secondary Content: The Linguistics of Side Issues*. Brill.

Westera, M., D. Goodhue, & C. Gussenhoven. 2020. Meanings of tones and tunes. In *The Oxford Handbook of Language Prosody*, 443–453. Oxford University Press.