On a concessive reading of the rise-fall-rise contour: contextual and semantic factors

Alexander Göbel & Michael Wagner*

Abstract. This paper presents three auditory rating experiments on the rise-fall-rise contour (RFR). Experiment 1 provides experimental evidence that the RFR makes disagreeing with a prior statement more natural than neutral intonation would. Additionally, the data show that the RFR exhibits a valence asymmetry, noted by Göbel (2019): the amelioration of a disagreement is greater when the RFR is used in a positive reply to a negative statement than in a negative reply to a positive statement. Experiments 2 and 3 investigate factors contributing to this asymmetry, showing that it disappears in replies to questions and is weakened when the reply contains an additive particle. Based on these results, we argue that the RFR has a scalar meaning, following Göbel (2019), with the relevant scale being contextually determined and resulting in an ambiguity resembling the Focus-particle *at least*.
Keywords. intonation; focus-particles; alternatives

1. Introduction. This paper addresses the question how intonation affects and conveys linguistic meaning through the lens of intonational contours, or "tunes", specifically, with a case-study on the so-called rise-fall-rise contour (RFR). The RFR has been characterized as conveying a sense of uncertainty or incompleteness on the speaker's part (Ward & Hirschberg 1985).[1] In apparent opposition to this characterization, the RFR can also be used more antagonistically, as in the naturally occuring example in (1). Here, Dewey's reply seems to challenge Hal's prior statement, but without straightforwardly contradicting it.

(1) *Hal*: I mean, you can't just break into a zoo, role a couple of elevens and suddenly become the dean of a university.
Dewey: *I* did... (*Malcolm in the Middle*: S7, E20; AUDIO)

The goal of this paper is to substantiate the tension between uses of the RFR as in (1) and its prior characterization empirically and explore relevant factors by presenting three auditory rating studies. We argue that there are in fact distinct uses of the RFR that resemble an ambiguity of the English Focus-particle *at least*, and support the proposal by Göbel (2019) that the RFR indicates the presence of a higher alternative on a variable scale.

The structure of the paper is as follows. Section 2 discusses details of prior accounts of the RFR and introduces the idea of a parallel between the RFR and *at least*. Section 3 presents the three experiments and Section 4 provides discussion of the results.

2. Background.

*We want to thank Massimo Lipari and Emma Nguyen for providing recordings for the studies, Andrea Beltrama, Chris Potts, Maribel Romero, and the audience at ELM2 for feedback on the project, as well as a Feodor-Lynen Fellowship of the Humboldt-Foundation to the first author and an SERC Discovery Grant to the second author for funding. Authors: Alexander Göbel, McGill University (alexander.gobel@mcgill.ca) & Michael Wagner, McGill University (chael@mcgill.ca).

[1] We will focus here on the meaning contribution of the RFR and leave equally important prosodic issues aside.

2.1. PRIOR ACCOUNTS. The seminal account by Ward & Hirschberg (1985) is primarily concerned with cases as in (2) where the RFR intuitively is used as a hedge that avoids a more definitive answer to the question while still intending to be relevant. Ward & Hirschberg's proposal centers the use of the RFR around speaker uncertainty in relation to a scale and its scalar values, either regarding the appropriateness of evoking such a scale, the particular choice of scale, or the choice of a newly added value on a given scale. Applied to (2), we could then say that there is uncertainty about "the relative positions of *Missouri* and *the Mississippi* on a geographical scale ordered from east to west" (W&H: 767).

(2) A: Have you ever been West of the Mississippi?
B: I've been to Missouri... (WARD & HIRSCHBERG 1985, (62))

In more recent work, the intuition that the RFR conveys uncertainty has been implemented in terms of incompleteness by Constant (2012) and Wagner (2012). For Constant, the RFR resembles Focus-particles like *only* in that it quantifies over alternatives and conveys that all assertable alternatives (= those alternatives whose truth-value is not yet known) cannot be safely claimed, formalized as in (3) as a conventional implicature.[2] Wagner proposes a similar analysis, with the difference that (i) relevant alternatives are speech acts rather than propositions and (ii) that there is an alternative speech act that could have been made, rather than focusing on what alternatives are not claimable, see (4).[3]

(3) $[\![\text{RFR } \phi]\!]^{ci} = \forall p \in [\![\phi]\!]^f$ s.t. p is assertable in C: the speaker cannot safely claim p.

(4) $[\![\text{RFR}]\!] = \lambda S.\ \exists S'\ in\ [\![S]\!]^g{}_a,\ S \nrightarrow S'$ and performing S' might be justified: S

These two accounts capture the distribution of the RFR in quantifier scales as in (5), where the RFR is felicitous with the middle value *some* but infelicitous with the extreme values *none* or *all*. *None* is odd because it rules out all higher alternatives such that the RFR applies vacuously since there are no alternatives left to be quantified over (Constant) or there is no alternative speech act to be made (Wagner), and *all* is odd for similar reasons since another alternative such as *some* is already entailed by it. In contrast, *some* leaves alternatives open that can then be marked as not being claimable or as the alternative speech act that was not made, hence the characterization of the RFR as conveying incompleteness (see also Goodhue et al. 2016 for experimental evidence).

(5) A: Did you feed the cats?
B: I fed { #none / some / #all } of them... (with RFR)

A different approach comes from Westera (2019) (see also Westera 2017), who focuses on the role of the RFR in the discourse structure, specifically the Question Under Discussion (QUD, Roberts 2012). According to Westera, the RFR indicates that a conversational maxim in relation to the main QUD has been suspended and instead a secondary QUD is being addressed. A potential benefit of this view is that it may allow including cases of contrastive topic, that share a similar prosodic profile, into a unified theory. However, to what extent such a unification is warranted is an open

[2] Constant (2012) also discusses the role of the RFR in scope ambiguities like *all...not*, which we put aside here, but see Syrett et al. (2014) for experimental data on this issue.

[3] Another difference regards the association with Focus, which is obligatory in Constant's theory but optional in Wagner's. We will come back to this issue in the General Discussion.

question.

2.2. GÖBEL (2019). The main approach that this paper builds upon comes from Göbel (2019). The central data point provided there is the observation that the RFR shows an asymmetry in replies to statements that contrast in the valence attributed to what is being discussed: while the RFR is natural as a response to A's negative characterization of Dexter in (6a) when descriptively opposing the previous statement with a positive characterization, the reverse - a negative reply to a positive statement - is markedly odd. We will dub this contrast in acceptability the "valence asymmetry".

(6) **Valence Asymmetry**

a. A: Dexter is such a horrible person. - B: He gives to charity... (AUDIO)

b. B: Dexter is such a great person. - A: ?? He murders people... (AUDIO)

Notably, this data point is unexpected on the accounts discussed in the previous subsection. To evaluate Ward & Hirschberg (1985), let's assume that the scale under consideration is a "goodness" scale, and what the replies convey is uncertainty about where to locate Dexter in virtue of his actions. It is unclear why (6a) would be more appropriate in this case than (8b) given that knowing someone gives to charity is compatible with being uncertain whether they are a horrible person, and knowing that someone murders people is compatible with being uncertain whether they are a great person.[4] For Constant (2012) and Wagner (2012), it is equally unclear why (6a) should leave alternatives open but not (6b), and for Westera (2019) there is no obvious difference in terms of the secondary QUDs (6a) and (6b) give rise to either.

The solution Göbel (2019) proposes is that the RFR indicates that there is an alternative that ranks higher on a given scale, essentially borrowing from ideas present in prior accounts. Let's again imagine a goodness scale that goes from pure evil on one end to pure virtue on the other, illustrated in Figure 1.

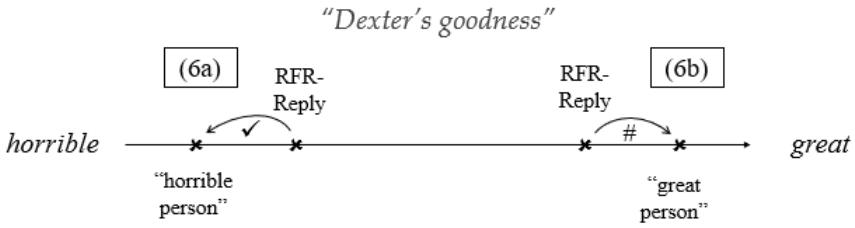

Figure 1: Illustration of scalar effects for (6).

In (6a), A's statement places Dexter toward the bottom of the scale. B's reply then suggests that Dexter is located higher. In contrast, in (6b) B's statement starts Dexter off toward the top and A's reply adds a negative counterpoint. As a consequence, the carrier utterance of the RFR does not contribute a higher alternative but a lower one and results in oddness.[5] The account additionally

[4] Admittedly, Ward & Hirschberg's account allows further possible interpretations of where the issue lies, but it is that flexibility that renders it less explanatory.

[5] A potential objection, however, would be that (6) is not a fair comparison: giving to charity is a less good indicator that someone is not a horrible person than murdering people is for indicating that someone is not a great person. One goal of the following experiment is to provide quantitative data that allows us to confirm or deny the reliability of the valence asymmetry.

predicts the contrast in (5): both *none* and *all* fail to leave higher alternatives open, while *some* leaves *all*.[6]

2.3. PARALLEL TO AT LEAST. Interestingly, the valence asymmetry in (6) is not unique to the RFR, but can also be observed with concessive (or evaluative) *at least* (Nakanishi & Rullmann 2009). The characteristic feature of concessive *at least* is that alternatives are ranked according to desirability, for instance in (7) conveying that feeding some of the cats is still better than feeding none of them.

(7) **At least** Cameron fed SOME of the cats (... it could've been less).

When using concessive at least in the same dialogues as (6) but with a falling contour, we see the same - intuitively even clearer - pattern: *at least* is felicitous with a positive reply (8a), while it is infelicitous with a negative reply (8b)

(8) a. A: Dexter is such a horrible person.
B: **At least** he gives to charity.
b. B: Dexter is such a great person.
A: # **At least** he murders people.

The parallel of the RFR with *at least* crucially does not end there. In addition to the concessive interpretation, *at least* also has an epistemic interpretation that has been tied to uncertainty (Geurts & Nouwen 2007). Taking (9) as illustration, the speaker only commits to Cameron having fed no less than some of the cats while leaving open the possibility that she fed all of them.

(9) Cameron fed **at least** SOME of the cats (... maybe even all of them).

This meaning of epistemic *at least* resembles prior accounts of the RFR that focus on its uncertainty or incompleteness component.[7] The resemblance goes as far as epistemic *at least* showing the same acceptability pattern with quantifiers as we saw in (5), namely being incompatible with extreme ends of a scale:

(10) A: Did Cameron feed the cats?
B: She fed **at least** { #none / some / #all } of them.

Based on this parallelism of the RFR with distinct interpretations of *at least*, we propose that the RFR is prone to a similar type of ambiguity. This view would allow us to reconcile theoretical accounts that emphasize the uncertainty and incompleteness component of the RFR with the valence asymmetry. Moreover, it would allow us to make sense of cases like (11), where the RFR is felicitous despite the truth of the higher alternative *ten* being known, showing that uncertainty or incompleteness are relevant but not sufficient. On the ambiguity view of the RFR, (11) would simply be a case of a concessive usage.

[6] Note that the case in (6) and the one in (5) differ on this view with respect to where the relevant alternatives come from: in (6), the alternative comes from the host utterance in relation to a previous one present in the discourse, while in (5), it is alternatives to the element that receives Focus that are relevant. We will come back to this issue later.

[7] A notable tension of this view comes from experimental results reported in de Marneffe & Tonhauser (2019), which show that the RFR strengthens implicatures when compared to neutral intonation. This finding is at odds with likening the RFR to an implicature suspension device like epistemic *at least*.

(11) A: Did you feed all ten cats?
B: I didn't feed all ten, but I fed NINE of them... (*with* RFR) (AUDIO)

However, an initial obstacle to this approach is the question what factors contribute to the ambiguity and determine the respective interpretation. We will explore this question in Experiments 2 and 3. Beforehand, Experiment 1 will be dedicated to substantiating the valence asymmetry empirically.

3. Experiments.

3.1. Experiment 1.

3.1.1. Materials & Design. The main goal of this experiment was to provide quantitative data on the valence asymmetry. To do so, we used short dialogues modeled after (6) that varied the "valence" of each utterance, shown in (12).

(12) a. *negative* + *positive* (= "*mismatch*")
A: The bike ride yesterday was really terrible, the weather was horrific.
B: We had a cocktail... (NEUTRAL, RFR)
b. *positive* + *negative* (= "*mismatch*")
A: The bike ride yesterday was really great, the weather was perfect.
B: We had an accident... (NEUTRAL, RFR)
c. *negative* + *negative* (= "*match*")
A: The bike ride yesterday was really terrible, the weather was horrific.
B: We had an accident...
d. *positive* + *positive* (= "*match*")
A: The bike ride yesterday was really great, the weather was perfect.
B: We had a cocktail...

The cases in (12a)-(12b) are labeled "*mismatch*" conditions since the valence of the context utterance and the target utterance differ. To complete the paradigm, we also added cases where valences where the same as a type of control, labeled "*match*" conditions.[8] Additionally, each target utterance was recorded in a *neutral* falling contour as a baseline and with an *rfr*. The design was thus a 2x2x2 (CONTEXT-VALENCE: *negative* vs *positive*, RESPONSE MATCH: *match* vs *mismatch*, INTONATION: *neutral* vs *rfr*).

We created 8 item sets like (12). Each item set was presented in all eight conditions but ordered in a way so that each participant saw each item set in all different conditions before seeing the item set again (within-design), for a total of 48 item trials.

3.1.2. Procedure. The experiment was implemented through prosodyExperimenter (https://github.com/prosodylab/prosodylabExperimenter) and ran online on Prolific.ac. Participants first saw a welcome screen, followed by a chance to adjust their volume and test their microphone, an online consent form, and a language background questionnaire. Afterwards, there was a test where participants were played three sounds and had to choose which one was the quietest, which required the use of headphones. For the main part of the experiment, participants were

[8] The manipulation could also be viewed in terms of whether speaker B agrees or disagrees with A's statement. We avoid using these labels here, however, since the extent to which a speaker's utterance is interpreted as a (dis-)agreement might be affected by intonation, such that focusing on the propositional content is more neutral.

auditorily presented with a dialogue and had to provide a naturalness rating on a scale from 1 (completely unnatural) to 6 (completely natural) based on how they thought the response sounded given the context. There were three practice trials after receiving instructions, followed by 48 stimuli. The experiment concluded with a chance to provide feedback. A test version of the experiment can be accessed at https://prosodylab.org/~agobel/conepi/conaddRatingBare/?SESSION_ID=test&mode=experiment.

3.1.3. PARTICIPANTS. 29 participants were recruited from Prolific.ac and paid $2.20 each. Five participants were excluded due to failing the headphone check, leaving 24 for data analysis.

3.1.4. RESULTS. *Coding & Data Analysis.* CONTEXT-VALENCE was sum coded given there was no basis for choosing one over the other as a baseline. INTONATION and RESPONSE MATCH, on the other hand, were dummy coded, with *match* and *neutral* as reference levels. Data were modeled with ordinal mixed effects with the maximal structure allowing convergence. Results and analysis files with more details on the full models of this and the following experiments, as well as experimental files and stimuli, can be found at the associated OSF repository: https://osf.io/t36qk

The average ratings by condition are shown in Figure 2. The most notable - although less interesting - pattern concerns higher ratings for *match* than *mismatch* conditions, reflected in a significant effect of RESPONSE MATCH ($z = 17.19$, $p < .001$***). Looking at only *match* conditions next, we see higher ratings for *neutral* intonation compared to the *rfr*, reflected in a significant effect of INTONATION ($z = -7.20$, $p < .001$***). However, when comparing the INTONATION difference for *match* with *mismatch*, the penalty for the *rfr* decreases, reflected in a significant interaction between INTONATION and RESPONSE MATCH ($z = 5.30$, $p < .001$***). Moreover, this amelioration of the penalty is greater with *negative* context sentences - and hence positive target sentences - than with *positive* context sentences, evidenced by a three-way interaction of INTONATION, RESPONSE MATCH, and CONTEXT-VALENCE ($z = 2.25$, $p < .05$*).

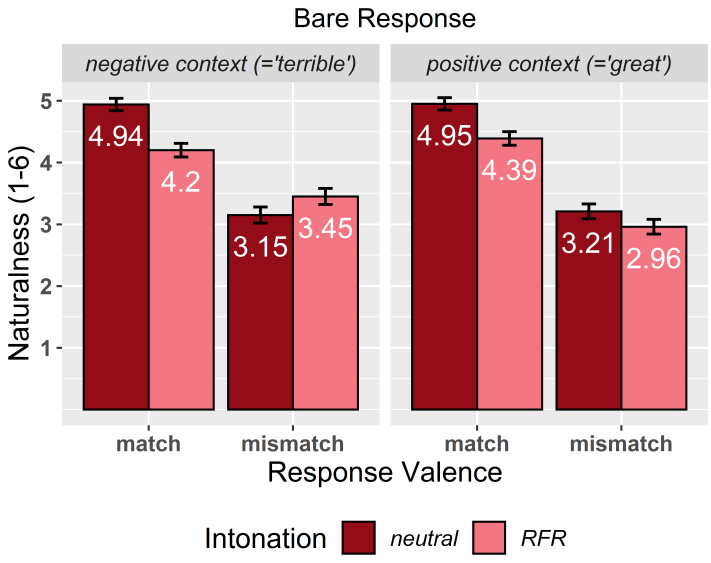

Figure 2: Naturalness ratings by condition, Experiment 1.

3.1.5. DISCUSSION. The experiment revealed several notable findings. First, participants strongly preferred dialogues where context utterance and target utterance had the same valence over dialogues that featured some opposition. While this was the numerically largest effect, it is also of least interest to this investigation such that we will put it aside. Second, the RFR ameliorated this mismatch penalty. Albeit novel, this finding is in line with prior accounts of the RFR insofar as expressing uncertainty or incompleteness may be viewed as more polite in an otherwise confrontational dialogue and hence improve ratings. Finally and most crucially, the extent to which the RFR ameliorated the penalty was larger in dialogues where the context utterance was negative and the carrier utterance of the RFR contributed a positive counterpoint compared to dialogues with a positive starting statement and a negative counterpoint. This last pattern is evidence for the valence asymmetry observed in Göbel (2019) and the argument made in this paper, namely that there is a concessive reading of the RFR alike to *at least*.

A follow-up question to this argument, mentioned in Section 2.3, is what makes this reading of the RFR available. A relevant hint in this regard comes from the fact that the prior literature has mostly focused on uses of the RFR in replies to questions or out of context, whereas Göbel (2019) and Experiment 1 investigated replies to (value) statements. A possible contrast between questions and statements is furthermore in line with intuitions and may thus be expected to contribute to the choice between an "epistemic" and a "concessive" reading of the RFR. The next experiment examines this connection.

3.2. EXPERIMENT 2.

3.2.1. MATERIALS & DESIGN. The goal of this experiment was to test whether replies to questions exhibit the same valence asymmetry we observed for replies to statements in Experiment 1. All design aspects of this experiment were identical to Experiment 1, with the sole exception that context utterances were changed from statements to questions, as in (13):

(13) a. *negative* + *positive* (= "*mismatch*")
A: Do you think yesterday's bike ride was terrible?
B: We had a cocktail...

b. *positive* + *negative* (= "*mismatch*")
A: Do you think yesterday's bike ride was great?
B: We had an accident...

c. *negative* + *negative* (= "*match*")
A: Do you think yesterday's bike ride was terrible?
B: We had an accident...

d. *positive* + *positive* (= "*match*")
A: Do you think yesterday's bike ride was great?
B: We had a cocktail...

3.2.2. PROCEDURE. The procedure was identical to Experiment 1. A test version of the experiment can be found here: https://prosodylab.org/~agobel/conepi/conaddRatingQ/?SESSION_ID=test&mode=experiment

3.2.3. PARTICIPANTS. 30 participants were recruited from Prolific.ac and paid $2.20 each. Six participants were excluded due to failing the headphone check, leaving 24 for data analysis.

3.2.4. RESULTS. *Coding & Data Analysis* were the same as Experiment 1 (see Section 3.1.4).

The average ratings by condition are shown in Figure 3. The first thing to note is that we again observe a mismatch penalty, with higher ratings for *match* than for *mismatch*, reflected in a significant RESPONSE MATCH effect (z = -8.89, p $<$ .001***). Additionally, ratings for *positive* contexts are lower than *negative* contexts in *match* cases, reflected in a significant effect of CONTEXT-VALENCE (z = 4.81, p $<$ .001***), but higher in *mismatch* cases, which shows up as a significant interaction of CONTEXT-VALENCE and RESPONSE MATCH (z = -5.47, p $<$ .001***). The second notable pattern is that there is little to no difference between *neutral* intonation and *rfr* across the board, with the exception of the *match* conditions in *positive* contexts, resulting in a marginal effect of INTONATION (z = 1.82, p $<$ .1•), a significant interaction of INTONATION and CONTEXT-VALENCE (z = -2.13, p $<$ .05*), a significant interaction of INTONATION and RESPONSE MATCH (z = -1.96, p $<$ .05*), and a marginal three-way interaction between INTONATION, CONTEXT-VALENCE, and RESPONSE MATCH (z = 1.65, p $<$ .1•).

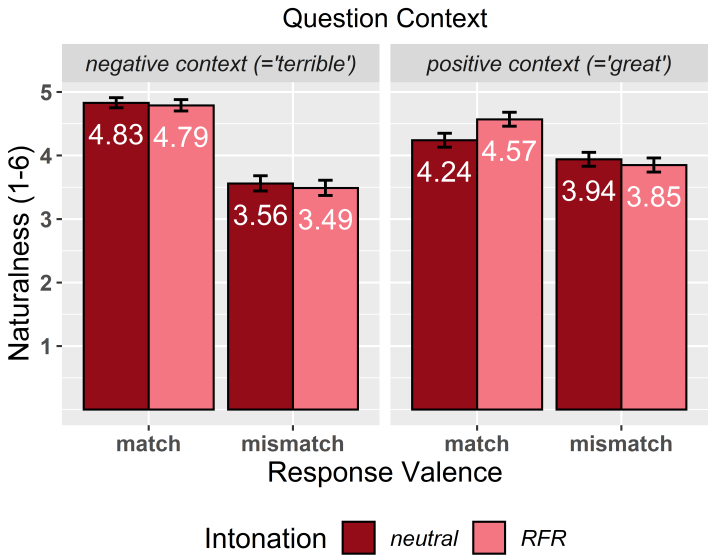

Figure 3: Naturalness ratings by condition, Experiment 2.

3.2.5. DISCUSSION. As in Experiment 1, participants considered dialogues that matched in valence more natural than those that mismatched. Moreover, this mismatch penalty was larger when the context question was negative than when it was positive. However, since this effect was largely independent of intonation, we will put it aside here.

More crucially, using questions - instead of statements as in Experiment 1 - seemed to render the difference between neutral intonation and the RFR almost non-existent. The only exception were higher ratings for the RFR in the match condition in positive contexts. Most importantly, there was no evidence for a valence asymmetry like we observed in Experiment 1. While the relevant three-way interaction was marginally significant, it went in the opposite direction to what we would have expected, namely with the RFR increasing the mismatch penalty in positive contexts rather than decreasing it less. The results thus provide support for the idea that using the RFR to reply to a question or to a statement affects what reading it receives.

Following Göbel (2019), we assume that the two readings can be modeled in terms of different scales that are determined pragmatically, for instance mediated by a QUD (e.g. Beaver & Clark 2008). For statements as in Experiment 1, the scale is evaluative, as discussed in sec 2.2. For questions as in this experiment, the relevant notion is truth, which can be modeled in terms of a scale of informativity. What is at stake in the question whether the bike ride was terrible or great then is to what extent the reply provides a definitive answer. Crucially, both positive and negative replies leave open "stronger", more definitive answers, such that the RFR is equally acceptable independently of valence, in contrast to what we saw in Experiment 1.

The following experiment explores a potential semantic factor for the interpretation of the RFR, namely the role of additive particles, based on the intuition that adding such a particle to the RFR utterance weakens the valence asymmetry.

3.3. Experiment 3.

3.3.1. Materials & Design. The goal of this experiment was to test whether an additive particle like *also* would neutralize or at least weaken the valence asymmetry observed in Experiment 1. Design and materials were identical to Experiment 1, except that the target utterances now contained *also*, as in (14):

(14) a. *negative* + *positive* (= "*mismatch*")
A: The bike ride yesterday was really terrible, the weather was horrific.
B: We also had a cocktail... (NEUTRAL, RFR)

b. *positive* + *negative* (= "*mismatch*")
A: The bike ride yesterday was really great, the weather was perfect.
B: We also had an accident... (NEUTRAL, RFR)

c. *negative* + *negative* (= "*match*")
A: The bike ride yesterday was really terrible, the weather was horrific.
B: We also had an accident...

d. *positive* + *positive* (= "*match*")
A: The bike ride yesterday was really great, the weather was perfect.
B: We also had a cocktail...

3.3.2. Procedure. The procedure was identical to Experiments 1 and 2. A test version of the experiment can be accessed here: https://prosodylab.org/~agobel/conepi/conaddRating/?SESSION_ID=test&mode=experiment.

3.3.3. Participants. We recruited 30 participants from Prolific.ac, who received $1.60 each. Six participants were excluded due to failing the headphone check, leaving 24 for data analysis.

3.3.4. Results. *Coding & Data Analysis* were the same as Experiments 1 and 2 (see Section 3.1.4).

The average ratings by condition are shown in Figure 4. The results pattern largely resembles that of Experiment 1: the biggest effect is a penalty for *mismatch* conditions compared to *match* conditions ($z = -15.13$, $p < .001$***), *neutral* intonation is rated better than *rfr* in *match* conditions ($z = -5.26$, $p < .001$***), but *rfr* makes *mismatch* cases less bad than *neutral* intonation ($z = 5.05$, $p < .001$***). However, in contrast to Experiment 1, the extent to which the mismatch penalty

amelioration of the RFR is greater in *negative* contexts than *positive* ones is less clear here, and in fact not supported by a significant three-way interaction (z = 0.63, p = .53).

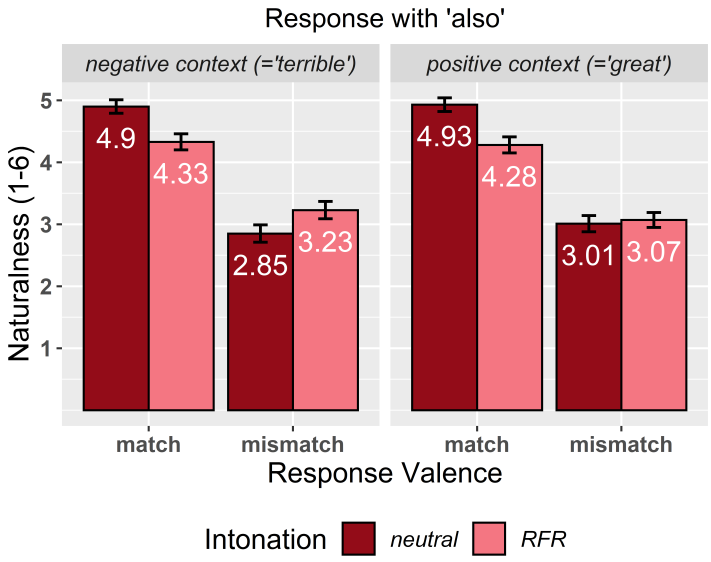

Figure 4: Naturalness ratings by condition, Experiment 3.

3.3.5. DISCUSSION. The results provided some support for the intuition that an additive particle weakens the valence asymmetry of the RFR: even though both the mismatch penalty and the amelioration of it by the RFR were comparable to Experiment 1 where *also* was absent, there was no statistical evidence that the amelioration differed between the two contexts. However, the numerical trend was still in line with a valence asymmetry. Moreover, the relevant comparison here is between the presence of an effect in Experiment 1 and the absence of the same effect in this experiment, rather than based on an effect within the same experiment, and should thus be taken as tentative evidence only. A proper confirmation that additive particles weaken the valence asymmetry will therefore be left for future research.

We may nonetheless wonder why additive particles would have such an effect. The potential explanation we want to put forward here is that the way the reading of the RFR gets determined interacts with the meaning of *also*. Additive particles are commonly analyzed as presupposing the truth of an alternative proposition (e.g. Heim 1992). In doing so, the interpretation of the RFR may become biased toward an "epistemic" reading that is concerned with informativity where valence is irrelevant and hence the valence asymmetry gets weakened. This approach would point toward an interesting issue of compositionality between two types of meanings - Focus-particles and intonational contours - that have been treated - including here - as making similar contributions based on relations to alternatives. How to spell out this connection in detail, however, goes beyond the scope of this paper and will be left for future work.

4. General Discussion. This paper presented data from three auditory rating experiments on the RFR. Experiment 1 substantiated a contrast between the acceptability of the RFR in positive replies to negative statements and negative replies to positive statements, as argued by Göbel (2019),

which we dubbed the valence asymmetry. We interpreted this finding as evidence for a "concessive" reading of the RFR, resembling concessive *at least*. Experiment 2 investigated replies to questions rather than statements and showed that the valence asymmetry disappears in this case. We took this change to point toward questions inducing a distinct, "epistemic" reading, again borrowing from a parallel with *at least*. On this view, the two readings can be unified by assuming that the RFR indicates the presence of a higher alternative but on different scales that are pragmatically determined. Finally, Experiment 3 explored the intuition that the inclusion of an additive particle like *also* weakens the valence asymmetry, for which we found tentative evidence. If confirmed, this finding raises interesting questions about how the meaning of intonational contours interacts compositionally with different parts of its carrier utterance.

Apart from this issue, there are two others we want to briefly discuss here. The first concerns the role of alternatives in the advocated theory. As mentioned earlier, the two readings are not taken to only differ in the type of scale involved but also in how they generate alternatives. For the "concessive" reading from Experiment 1, the relevant higher alternative is contributed by the host utterance of the RFR, with a prior statement being the proposition that the alternative is higher than. For the "epistemic" reading from Experiment 2, the higher alternative is an alternative to the host utterance itself, similar to how the assumed alternative to *some* in the quantifier cases (5) from Section 2.1 is *all*. A prediction of this view then is that the RFR should in principle be compatible with extreme scale values as long as it contributes an alternative that is higher than some previously mentioned one and receives a concessive reading.[9] Future research will have to show whether this prediction is borne out.

The second issue is about the meaning of intonation at large. Here, we took the RFR to contribute its meaning holistically rather than attempting to attribute associate smaller meaning components with the individuals parts of the contour, but we can ask what such a decompositional account may look like. The RFR is prosodically analyzed as, using ToBI annotation, consisting of an L*+H pitch accent, a L- phrasal accent, and a H% boundary tone. We will put the phrasal accent aside here and focus on the pitch accent and the boundary tone. According to Pierrehumbert & Hirschberg (1990), the L*+H accent evokes a scalar meaning, which is consistent with the view of the RFR adopted here. Regarding the high boundary tone, there are a number of recent accounts of rising declaratives to draw from. Rudin (2022) proposes that the characteristic feature of rising intonation is that they lack speaker commitment, which seems incompatible with the valence asymmetry given that for both types of replies the speaker is intuitively committed to the truth of the proposition. Jeong (2018) argues that rising intonation opens a meta-linguistic issue, which again seems to lack a clear path for explaining the valence asymmetry per se, but may capture the intuitive sense of indirectness and the resulting politeness effect. We aim to address the question about the contribution of the final rise in future work.

References

Beaver, David & Brady Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Oxford: Blackwell.

[9]This difference is furthermore reminiscent of and possibly connected to the issue between Constant (2012) and Wagner (2012) regarding whether the RFR obligatorily associates with Focus or not.

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

Geurts, Bart & Rick Nouwen. 2007. 'at least' et al.: the semantics of scalar modifiers. *Language* 83. 533–559. https://doi.org/10.1353/lan.2007.0115.

Goodhue, Daniel, Lyana Harrison, Y. T. Clémentine Su & Michael Wagner. 2016. Toward a bestiary of english intonational contours. In Brandon Prickett & Christopher Hammerly (eds.), *Proceedings of the North East Linguistics Society 46*, 311–320.

Göbel, Alexander. 2019. Additives pitching in: L*+H signals ordered focus alternatives. *Proceedings of Semantics and Linguistic Theory (SALT) XXIX* 1–11.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9. 183–221.

Jeong, Sunwoo. 2018. Intonation and sentence-type conventions: Two types of rising declaratives. *Journal of Semantics* 35. 305–356.

de Marneffe, Marie-Catherine & Judith Tonhauser. 2019. Inferring meaning from indirect answers to polar questions: The contribution of the rise-fall-rise contour. In Edgar Onea, Malte Zimmermann & Klaus von Heusinger (eds.), *Questions in discourse*, 132–163. Leiden: Brill.

Nakanishi, Kimiko & Hotze Rullmann. 2009. Epistemic and concessive interpretation of *at least*. Talk presented at Canadian Linguistics Association. https://linguistics.sites.olt.ubc.ca/files/2018/03/2009.Nakanishi_Rullmann.CLA_-1.pdf.

Pierrehumbert, Janet B. & Julia Hirschberg. 1990. The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan & M. E. Pollack (eds.), *Intensions in communication*, 271–311. Cambridge, MA: MIT Press.

Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5. 1–69. https://doi.org/10.3765/sp.5.6. Earlier version appeared in OSU Working Papers in Linguistics 49 in 1996.

Rudin, Deniz. 2022. Intonational commitments. *Journal of Semantics* 39. 339–383.

Syrett, Kristen, Georgia Simon & Kirsten Nisula. 2014. Prosodic disambiguation of scopally ambiguous quantificational sentences in a discourse context. *Journal of Linguistics* 50. 453–493.

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

Ward, Gregory & Julia Hirschberg. 1985. Implicating uncertainty: the pragmatics of fall-rise intonation. *Language* 61. 747–776.

Westera, Matthijs. 2017. *Exhaustivity and intonation: a unified theory*: University of Amsterdam dissertation.

Westera, Matthijs. 2019. Rise-fall-rise as a marker of secondary QUDs. In Daniel Gutzmann & Katharina Turgay (eds.), *Secondary content: the linguistics of side issues*, Leiden: Brill.