Relating Scalar Inference and Alternative Activation: A view from the Rise-Fall-Rise Tune in American English

The rise-fall-rise (RFR) tune in American English is notorious for its varied and often conflicting description in both its pragmatic function and its phonological description. Phonological theory (Pierrehumbert, 1980) predicts 3 RFR-shaped tunes that differ in pitch accent (monotonal H* or bitonal L+H* or L*+H), yet it is unclear whether reported variation in the semantic/pragmatic function of RFR is due to differences among studies in the intonational materials used, or whether some (or perhaps all) accounts might be unified under a broad class of RFR intonational patterns. A common thread among accounts relates RFR to higher alternatives; i.a. conveying uncertainty (Ward and Hirschberg, 1985), unclaimability (Constant, 2012), or salience (Göbel, 2019) of some higher alternative. These accounts make different predictions for RFR's effect on scalar inference calculation (SI, e.g., *some→ some but not all*). Experimental work has used SI as a probe to adjudicate among accounts of RFR, finding higher rates of SI calculation when RFR is used (de Marneffe and Tonhauser, 2019; Göbel and Ronai, 2023, though cf. Buccola and Goodhue, 2023); however, these studies tested only a single RFR tune with non-specific phonetic description. Thus, there lacks a study comparing the interpretation of the various RFR-shaped tunes. Finally, there is a processing question: results from cross-modal priming show that some but not all rising pitch accents modulate the activation of focus alternatives in processing (i.a., Husband and Ferreira 2016). Since alternatives in SI and focus have been claimed to be related both in semantic/pragmatic theory (i.a., Fox & Katzir, 2011) and psycholinguistics (i.a., Gotzner & Romoli, 2022), differences in the interpretations of RFR-shaped tunes may be reflected in processing, as indexed by alternative activation. We present a systematic investigation of 3 RFR-shaped tunes. Using both SI judgment and cross-modal lexical decision tasks, we test 1) differences among RFR-shaped tunes in their interpretation, and 2) whether any such effects are mirrored in processing. **Materials:** We wrote polar question+indirect answer dialogues for 72 different adjective pairs that form a scale, e.g., <*tough, impossible*> Q: *I haven't gone running since before the pandemic, do you think I could do a half marathon?* A: *That distance would be **tough***. In a text-only norming task, undergraduate participants (n=48) read the dialogues, provided acceptability ratings, and answered questions such as "Would you conclude that that distance would **not** be **impossible**?", where a "Yes" (as compared to "No") response means that SI was calculated. We found that the dialogues were overall acceptable compared to incongruent fillers and we replicated previous findings that SI calculation rates vary across scales (i.a., van Tiel et al., 2016). We chose 64 items to record in 6 intonation conditions with one of 3 pitch accents ("neutral" H* and focus-marking L+H* and L*+H) and one of 2 edge-tones (fall, L-L%, or fall-rise L-H%) and standardized the pitch contours using pitch resynthesis in Praat.

Exp 1: These auditory materials were used in a follow-up SI task. Online participants (Prolific, n=83) listened to a dialogue in one of 6 intonation conditions then answered questions like "Would you conclude […] not be impossible?" If RFR conveys uncertainty about a higher alternative, and not belief in its negation, then we predict lower SI rates for RFR compared to falls. But if RFR instead functions more broadly to mark the salience of higher alternatives, we predict higher SI rates for RFR, as salient alternatives are available for SI calculation. Such effects could potentially be seen with just **one** RFR or with **any** RFR-shaped tune. **Results:** A Bayesian logistic mixed effects model shows a main effect of edge-tone: all RFR-shaped tunes yielded higher SI rates compared to falls (posterior probability of direction=100%). We find slight gradience between the pitch accents within the RFR tunes, with an average SI rate ranking of L*+H > L+H* > H*, which is reversed for falls (Fig. 1).

Exp 2: In a cross-modal lexical decision task, participants listened to the recorded dialogues (*…tough)* and, after a 750ms delay, wereshown the higher alternative (*impossible*). Participants then judged whether the displayed string was a word or non-word. We measure the reaction time (RT) of participants' judgments and predict that if RFR invokes higher alternatives (as shown in Exp 1), then *impossible* should be facilitated (=faster RTs). To control for activation arising from semantic similarity, adjectival scales are also tested in the opposite prime-target order, where participants listen to *…impossible* and judge *tough.* The predicted RFR effect on the activation of higher alternatives (e.g. *impossible*) should be greater than (1) the activation of *tough* when *impossible* is uttered with RFR and (2) the activation of *impossible* when *tough* is uttered with other tunes. This task is administered in the laboratory using low-latency hardware in a sound-attenuated booth, as well as online, where the hardware and environment of the participant cannot be as easily controlled. Preliminary results from in-person data collection (Undergraduates, n=46/target 60) (Fig.2) suggest a main effect of displayed alternative such that RT is faster for higher alternatives after accounting for word frequency and length (p.d.=98%). Moreover, H*L-H% shows evidence of an interaction, yielding additional facilitation for higher alternatives (p.d.=91.9%). We do not find evidence of such an effect with the other RFR-shaped tunes (p.d.<85%) nor the falls (p.d.<65%). Online participants (Prolific, n=60) show longer and more varied RTs with no notable pattern of facilitation across intonation conditions, suggesting that this in-person effect is subtle and not robust to noise arising from unconstrained environmental factors.

Discussion: The findings of Exp. 1 replicate prior work and are most compatible with accounts of RFR that do not invoke uncertainty. But the combined findings from the two tasks present a puzzle: in Exp. 1, while all RFR-shaped tunes increase SI rates compared to falls, L*+HL-H% increases the likelihood of SI most strongly. Since SI arises via retrieval and negation of a higher alternative, we expect the processing signature of this tune to show stronger facilitation in lexical decision. Yet in the lexical decision task, only H*L-H% provides evidence of facilitation. We discuss a possible account of this pattern in terms of pitch range, following Ward & Hirschberg (1992): expanded pitch range in the bitonal RFRs may invite competing inferences, e.g., related to speaker arousal. In the SI task, participants might be more willing to accommodate the SI-enriched interpretation because it is explicitly probed via the question in the trial. But this task effect is not present in priming, and the competing inferences may mask potential facilitation.

Implications: By controlling the intonational variation in the auditory materials and the experimental setting, we provide novel psycholinguistic evidence of the relationship between RFR-shaped tunes and scalar alternatives. We find a distinction between lower-scaled/monotonal RFR (H*) and higher-scaled/bitonal RFR (L+H*/L*+H), but overall, all RFR tunes behave differently from falls. The variation among RFR-shaped tunes in our experiments is emblematic of the variation seen in prior work, suggesting within-category variation may reflect more particularized inferences beyond RFR's conventional connection with higher alternatives.
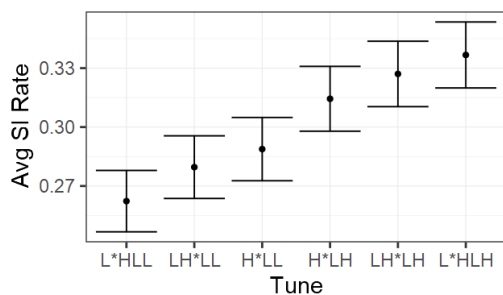

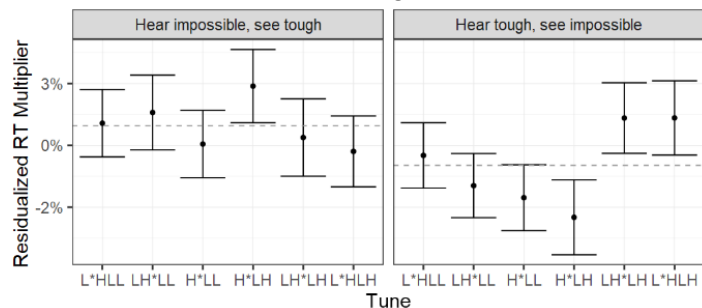*Fig. 2: Average SI rates for each tune in Exp. 1, _LL=Fall & _LH=RFR*

Fig. 1: Residual speedup (-x%) or slowdown (+x%) in RT controlled for log word frequency, log length, and block.