Rising pitch, continuation, and the hierarchical structure of discourse

Joseph Tyler

The meaningful contribution of terminal rising pitch has received a fair amount of scholarly attention, discussed for its ability to create questioning force on declarative syntax (Gunlogson 2008), as part of listing intonation (Ladd 2008), as well as indicating discourse relationships (Jasinskaja 2010, Nilsenová 2006, Pierrehumbert and Hirschberg 1990). Sometimes it is discussed simply as a final rise, like in the above work, while other times as part of a rise-fall-rise contour, sometimes called contrastive pitch accent (Büring 2003, Constant 2012). A common interpretation of its meaning is that it conveys incompleteness, more-to-come, continuation or is 'forward-looking' (Bolinger 1989, Hirschberg 2008, Pierrehumbert and Hirschberg 1990). Recent experimental results contribute to this discussion, showing a rise can bias towards the coordinating interpretation of a coordination/subordination discourse ambiguity (Tyler 2012). Because both interpretations of the ambiguity involve continuation, the rise is signaling not just that you continue but *how* you continue. In this paper, I will briefly present these results and then integrate them into a unified account of the contribution of rising pitch, which I see as a signal of incompleteness with respect to the current hierarchical level of the discourse.

Experimental results in Tyler (2012) show that prosody can bias the interpretation of ambiguously structured discourse. The experiment involved ambiguous discourses like (1):

(1) I sat in on a history class. I read about housing prices. And I watched a cool documentary.

The discourse in (1) could be interpreted such that the narrator read about housing prices and watched a cool documentary in history class (Subord interpretation) or separate from history class (Coord interpretation). In the Subord interpretation, sentences 2 and 3 provide further information, i.e., elaborate, the event described in sentence 1. In the Coord interpretation, each sentence describes a separate event.

An experiment was designed to test whether prosody could bias listeners' interpretation of discourses like (1). Specifically, a contrast of rising vs. falling terminal pitch on the first sentence of discourses like (1) was constructed for its ability to bias listeners towards Subord and Coord interpretations.

First, a set of over 100 discourses with structural ambiguity like (1) were normed for how often participants interpreted them as Coord or Subord. Participants in the norming study were asked to choose which one of three possible interpretations they thought the narrator Sally meant:

Did Sally read about housing prices and watch a cool documentary in history class? (Subord)
Did Sally read about housing prices and watch a cool documentary separate from history class? (Coord)
Did Sally mean something other than these two interpretations?

Forty-seven participants, recruited through Amazon Mechanical Turk, took part in the norming study.[1] Discourses that got a high proportion of "other" interpretations were excluded. The 48 discourses most ambiguous between Coord and Subord were selected for inclusion in the experiment.

For the experiment, participants listened to 48 discourses structured like (1), 24 with rising and 24 with falling pitch at the end of the first sentence. Every sentence of every discourse was read aloud in a carrier context by a female native speaker of American English in the sound lab at the University of Michigan. The carrier context involved embedding each target sentence between the same two sentences, and randomizing the order of the target sentences, e.g., (2).

(2) I'm going to read a sentence. I read about housing prices. I just read a sentence.

---

[1]For more detail on the norming, see (Tyler, 2012).

All target sentences were then spliced out and concatenated with intervening pauses to create the discourses. To create the experimental contrast, sentence 1 (e.g., "I sat in on a history class") was synthetically manipulated to create a rise or fall from the last stressed syllable to the end. Sentence 2 final pitch was synthetically flattened. Pitch was manipulated within the Pitch Manipulation object in Praat. The only contrast between the two conditions was the rise vs. fall at the end of sentence 1.
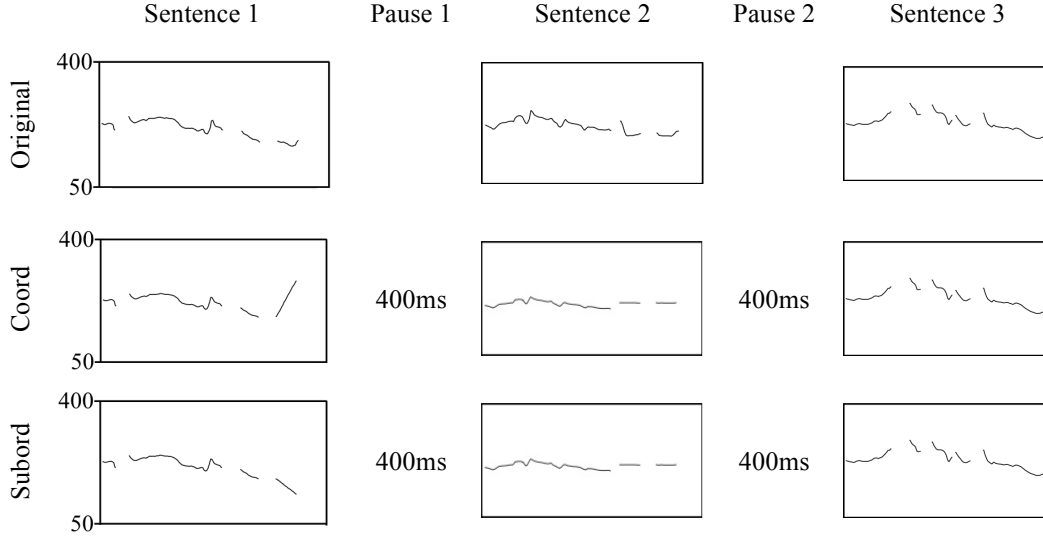

Table 1: Pitch contours for the three sentences in (1).

Table 1 shows pitch contours of the original productions on the top row, the Coord prosody manipulation in the second row and the Subord prosody manipulation in the third row. The contours show a linear rise on sentence 1 for Coord and linear fall for Subord, as well as flat pitch at the end of sentence 2 for both Coord and Subord conditions. Inter-sentential pause durations were held constant at 400ms.

Fifty-eight participants were recruited through Amazon Mechanical Turk for the experiment. The discourses were counterbalanced such that each participant heard each discourse only once, but that overall each discourse was heard equally often in rise and fall conditions. After hearing the discourse, participants were asked an interpretation question (e.g., Did Sally [the supposed speaker] read about housing prices and watch a cool documentary in history class?). Their answer was coded for whether they interpreted the discourse as Subord or Coord.

The result of most interest was whether the pitch manipulation had an impact on discourse interpretation. Results were modeled statistically with a mixed effects model with binomial outcome fit using the `lmer` function in the R package `lme4`. The dependent variable was matching vs. mis-matching interpretations given the prosodic condition, where match was defined as a rise resulting in a Coord interpretation and a fall resulting in a Subord interpretation. The model included random effects for subject and item, as recommended elsewhere (Barr, Levy, Scheepers, and Tily 2013, Jaeger 2008), in order to control for inter-subject and inter-item variation. Results show a difference in the likelihood of match vs. mismatch of prosody and interpretation (ß=.186, SE=.067, z=2.782, p=.005), showing that the pitch manipulation did affect discourse interpretation. Rises led to 52% Coord interpretations and falls led to 43% Coord interpretations, showing that the rising pitch on sentence 1 biased participants towards more Coord interpretations relative to a fall. Moreover, whether a participant's interpretation matched the prosody also predicted their confidence (ß=2.037, SE=.597, t=3.41, p_{MCMC}=.001). Participants were more confident in their interpretation when they chose the Coord interpretation after hearing a rise or a Subord interpretation after hearing a fall, compared to the alternative combinations of Coord/fall or Subord/rise.

These results show that when confronted with a discourse that is ambiguous between a Coord and Subord interpretation, terminal rising pitch can bias towards the Coord interpretation. In both interpretations, the second sentence attaches to the first, but the difference is in the nature of that attachment. In the Subord interpretation, sentences 2 and 3 are embedded at a lower level of structure while in the Coord interpretation they are all at the same hierarchical level.[2] Therefore, an analysis of the meaning of rising pitch as continuation alone glosses over the differences in kinds of continuations that rising pitch can help disambiguate between. The goal then becomes to specify what kind of continuation rising pitch is conveying in contrast to falling pitch. As suggested by discussing the two interpretations as Coord and Subord, I propose that the kind of continuation is distinguished with respect to the structure of discourse. In addition to the discourse structuring effect described above, the rise seems to have similar discourse structuring properties as rises that create questioning force on declarative syntax. While detailed below, the general idea is that questions and discourse segments followed by coordinated discourse segments each make an incomplete contribution to a question under discussion (QUD). The effect of a question is to introduce a question under discussion that needs an answer. The effect of a coordinating relation is to indicate partial answerhood to some dominating question under discussion. Either way, the rise conveys incompleteness (question, partial answer) with respect to the discourse.

To flesh this proposal out more, I will convert the coordination/subordination ambiguity in (1) to a QUD structure (Büring 2003, Roberts 1996). The discourse ambiguity in (1) was modeled by Tyler as ambiguous between a Narration (coordination) structure and an Elaboration (subordination) structure using *Segmented Discourse Representation Theory* (SDRT) (Asher and Lascarides 2003). I propose the two meanings of (1) could be modeled as having the two QUD structures below, where (3) and (4) correspond to the Coord and Subord interpretations of (1) respectively.

(3) What did you do yesterday?
    a.   I sat in on a history class. RISE
    b.   I read about housing prices.
    c.   I watched a cool documentary.
(4) What did you do yesterday?
    a.   I sat in on a history class. FALL
    b.   What did you do in the history class?
        i.   I read about housing prices.
        ii.   I watched a cool documentary.

Both interpretations in (3) and (4) of the discourse in (1) are felicitous answers to the superdominant question *What did you do yesterday?*. The interpretations differ structurally by whether sentence 2 of (1) continues at the same hierarchical level ((3)b) or a different hierarchical level ((4)b.i). Sentence ((3)b) continues to answer the superdominant question of what you did yesterday while sentence ((4)b.i) answers a subquestion ((4)b) of what you did in history class. That is, ((3)a)-((3)c) all provide answers to the same question about what you did yesterday. By contrast, in (4) only ((4)a) directly answers the question about what you did yesterday, while ((4)b.i)-((4)b.ii) answer a subquestion about what you did in history class. What the results in Tyler (2012) show is that a rise/fall contrast at the end of the first sentence can help disambiguate between these two structures. Because a rise ending sentence 1 resulted in more offline interpretations like (3), it may be the case that the rise cues the listener that the following sentence is more likely to be at the same level of hierarchical structure (e.g., (3)b) relative to an alternative at a deeper level of hierarchical structure (e.g., (4)b.i). In this analysis, the rise then conveys incompleteness at the current level of hierarchical structure, not just incompleteness more generally. The intuition that a rise indicates continuation (leading it at times to be called a "continuation rise") is actually about continuation at the current level of hierarchical structure in the discourse. In the context of discourses like (1), the contribution of the rise as a kind of discourse incompleteness at the current level of hierarchical structure indicates partial answerhood to the superdominant question *What did you do yesterday?*.

---

[2]For more on how the Coord vs. Subord hierarchical contrast can be modeled, see (Tyler, 2012), which uses Segmented Discourse Representation Theory (Asher and Lascarides, 2003) to represent the contrast.

The intonation of lists also provides evidence in support of the analysis above. The interpretation in (3) views each sentence as denoting a list of separate events the speaker did. Work on the intonation of lists, however, has not generally looked at a list composed of whole sentences, focusing instead on lists of either noun phrases (NPs) or verb phrases (VPs). So, for example, Beckman and Pierrehumbert use the sentence fragment in (5) and Ladd gives the example in (6) where each of the listed elements is an NP. And Cauldwell and Hewings provide an example of listed VPs in (7).

(5) Blueberries, bayberries, raspberries, mulberries and boysenberries (Beckman and Pierrehumbert 1986:273)
(6) I need milk and eggs and butter and bread (Ladd 1980:183)
(7) John has got to buy some coffee, wash the floor, and wind the clock (Cauldwell and Hewings 1996:330)

The Coord interpretation of (1) shows a list that is composed of whole sentences. The motivation for looking at lists and listing intonation here, however, is that lists provide a more transparent case of elements at the same level of hierarchical structure. The elements in a list are all at a similar level of detail and share some similar relationship to the superordinate category defining the list. So, the elements in a list of berries (e.g., blueberries, bayberries, raspberries, mulberries) are all equally members of the set of berries. There is no direct hierarchical relationship between the berries themselves, and in this sense, they all equally answer the relevant question under discussion (e.g., *what berries did you eat?*).

One feature of lists is that the members can be re-arranged without changing meaning. That is, the sentence *I like apples, plums and pears* means the same thing as *I like pears, apples and plums*, where the re-ordering of the fruits does not change the propositional content of the utterance. Analogously, the sentences of the discourse in (1), when interpreted as a list like in (3), can be re-arranged without changing meaning. But re-arranging the sentences of (1) makes it difficult to get the Subord interpretation in (4). And while a list of fruits answers a question under discussion like *what fruits did you eat*, a list of activities answers a question under discussion like *what activities did you do*.

There is some work on the intonation of lists, and while there is little experimental or corpus work on the actual production of lists, there seems to be general agreement on what listing intonation is, at least in its canonical form. This canonical listing intonation is characterized as a series of rises on non-final members of the list and a fall on the last one (Cauldwell and Hewings 1996, Hirschberg 2008, Ladd 1980, Schubiger 1958). Ladd gives the example "I need milk and eggs and butter and bread," with a rise on milk, eggs, butter and a fall on bread, as an example of a sentence that would be produced with canonical listing intonation (1980:183-184). This suggests that when the discourse in (1) is interpreted as a list of separate activities, i.e., (3), it could be conveyed with the canonical set of pre-final rises and concluded with a fall. Moreover, so-called listing intonation could provide one way to help disambiguate between a list interpretation and an embedding interpretation. This discussion of listing intonation is intended to show that the structural relationship between members of a list as hierarchically equal, and the associated intonation of rises on non-final elements of the list, also seems to be operative at the level of discourse when each listed item is a whole sentence. Therefore, it is reasonable to consider the Coord interpretation of ambiguous discourses like (1) as having a list structure, and the sentences of those discourses as being hierarchically equal to one another in the QUD structure of the discourse. It is also reasonable to expect that listeners might draw on listing intonation in their interpretation of what the speaker meant upon uttering the discourse.

One of the most influential accounts of intonational meaning comes from Pierrehumbert and Hirschberg (1990), who attempt to provide a compositional analysis of the pitch accents and boundary tones of an autosegmental-metrical theory of intonational phonology. This approach treats each pitch accent type and boundary tone type, as captured in the ToBI (Silverman et al. 1992) transcription paradigm, as carrying its own distinct meaning that in combination with the others aggregates to the contribution of the intonation. This discussion will focus on the high boundary tone, written as H% in ToBI, because it tends to correspond to terminal rising pitch. Pierrehumbert and Hirschberg propose that boundary tones in general indicate something about

how one intonational phrase relates to another. They write: "boundary tones convey information about relationships among intonational phrases- in particular, about whether the current phrase is to be interpreted with particular respect to a succeeding phrase or not" (1990:287). A high boundary tone H% is suggested to introduce hierarchically lower discourse segments, e.g., through a relation like elaboration. This claim is motivated by the following example:

(8)
    a. The train leaves at seven
           H*   H*   H* L H%
    b. It'll be on track four
           H*   H* L L%

Using the terminology of the Grosz and Sidner (1986) model of discourse structure, Pierrehumbert and Hirschberg argue that "the satisfaction of the purpose S [the speaker] has in uttering ((8)b) contributes to the satisfaction of S's purpose in uttering ((8)a) by further elaboration" (1990:287). They posit ((8)a) is in a dominance relationship with ((8)b), because the goal of ((8)b) serves to partially fulfill the goal of ((8)a). On the surface, this account that H% can indicate an elaboration relation appears to contradict the results in Tyler (2012), which showed that rising pitch biased listeners away from an elaboration interpretation. I propose this apparent contradiction can be resolved by a re-analysis of the structure of (8), so that instead of being related hierarchically via elaboration, ((8)a) and ((8)b) actually together elaborate some implicit superordinate topic. One definition of an elaboration relation comes from an annotation manual (Reese, Denis, Asher, Baldridge, and Hunter 2007) for *Segmented Discourse Representation Theory* (SDRT) (Asher and Lascarides 2003) created for a project called DISCOR (Baldridge, Asher, and Hunter 2007) that involved developing an SDRT-annotated corpus of newspaper articles. In this manual, the authors write "Elaboration(α, β) holds when β provides further information about the eventuality introduced in α; for example, if the main eventuality of β is a sub-type or part of the eventuality mentioned in α" (Reese, et al. 2007:7). In other words, an elaborating discourse segment is one that provides more information about the eventuality in the discourse segment it elaborates. If we apply this definition of elaboration to (8), the claim that ((8)b) elaborates ((8)a) would mean that the train being on track four provides further information about the train leaving at seven. Sentence ((8)a) ("The train leaves at seven") indicates "time of departure," while sentence ((8)b) ("It'll be on track four") indicates "location of departure". It is unclear how location of departure could provide further information about time of departure. Rather, both of these seem to provide further information about, i.e., elaborate, the departure itself.

    Under this analysis, ((8)a) and ((8)b) elaborate an implicit topic about the train's departure. Within the terminology of a QUD structure, both ((8)a) and ((8)b) are answers to a question under discussion like *What do you know about the train to London?*. The consequence of this re-analysis is that both ((8)a) and ((8)b) are hierarchically equal in the discourse, just like all three sentences in the interpretation in (3). As a result, the effect of rising pitch in Tyler (2012) and the meaning of H% in Pierrehumbert and Hirschberg (1990) are both to mark incompleteness at the same level of hierarchical structure. Therefore, the continuation part of what is often called a continuation rise is continuation specifically at the same level of hierarchical structure.

    In contrast to this proposal of rising pitch as incompleteness with respect to discourse structure, Nilsenova (2006) provides an account of rising pitch as uncertainty. She argues rises are a kind of intonational adverb, analogous to a modal structure like "Ist might be that…". If we apply this to the discourse in (1), the version with and without a rise could be paraphrased as the following.

(9) Coord: It might be that I sat in on a history class. I read about housing prices. And I watched a cool documentary.
(10)   Subord: I sat in on a history class. I read about housing prices. And I watched a cool documentary.

This would mean that the speaker is uncertain about whether they sat in on a history class in (9). But the difference in interpretation found in Tyler (2012) is not with respect to epistemic uncer-

tainty, but with respect to the interpretation of the reading and watching events of (1) as happening during history class or not. At least for these data, it is hard to see how the uncertainty interpretation can account for the effect found in Tyler (2012).

Büring (2003) discusses how intonation that cues focus can affect the structure of discourse ("d-trees" in his terminology). He discusses two ways discourses can be infelicitous: incoherence (the discourse is "defective in content proper") and incongruence (the discourse is "defective in (intonational) form"). This suggests that particular discourse structures can be produced with incongruent intonation. In the context of discourses like (1), the mismatching cases of Subord interpretations with rising pitch and Coord interpretations with falling pitch could be termed incongruous. Büring cashes out his understanding of incongruence in terms of probabilistic weights, not categorical contrasts. This probabilistic explanation fits with the fact that the rise in Tyler (2012) biases interpretation, but not categorically. Further evidence of the incongruence of Subord interpretations after hearing a rise and Coord interpretations after hearing a fall can be seen in the higher confidence participants showed in matching, or "congruent", interpretations (Tyler 2012).

Jasinskaya (2010) discusses rising pitch as a trigger away from default interpretations. One of her defaults is topic continuity, stating that without external effects (e.g., from discourse markers, topic shifts, or intonational marking), speakers would prefer to continue on the same topic. This suggests that speakers would prefer the Subord interpretation of (1), because once the history class event is introduced they would rather continue on that topic. The norming study in Tyler (2012) suggests that topic continuity as a default may be problematic, as the 100 ambiguous discourses showed a continuum of preferred interpretation from more Coord to more Subord. This shows that lexical content, and thus world knowledge, can bias what is more default. On the other hand, Jasinskaya's claim that intonation can trigger interpretation away from a topic continuity default is in line with Tyler's experimental results. Tyler found rising pitch at the end of the first sentence of discourses like (1) biased listeners more towards a Coord (i.e., new topic) interpretation. In this way, her interpretation of the role of intonation parallels the results of Tyler (2012) and the account presented here.

References

Asher, Nicholas, and, Alex Lascarides. 2003. *Logics of Conversation*. Cambridge, UK: Cambridge University Press.

Baldridge, Jason, Nicholas Asher, and Julie Hunter. 2007. Annotation for and Robust Parsing of Discourse Structure on Unrestricted Texts. *Zeitschrift fur Sprachwissenschaft* 26: 213−239.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255−278.

Beckman, Mary, and Janet Pierrehumbert. 1986. Intonational Structure in Japanese and English. *Phonology Yearbook* 3: 15−70.

Bolinger, Dwight. 1989. *Intonation and its Uses: Melody in Grammar and Discourse*. Stanford, California: Stanford University Press.

Büring, Daniel. 2003. On D-Trees, Beans, and B-Accents. *Linguistics and Philosophy* 26: 511−545.

Cauldwell, Richard, and Martin Hewings. 1996. Intonation rules in ELT textbooks. *ELT Journal* 50: 327−334.

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

Grosz, Barbara, and Candace Sidner. 1986. Attention, Intentions, and the Structure of Discourse. *Computational Linguistics* 12: 175-204.

Gunlogson, Christine. 2008. A Question of Commitment. *Belgian Journal of Linguistics* 22: 101−136.

Hirschberg, Julia. 2008. Pragmatics and Intonation. In *The Handbook of Pragmatics*, ed. L. Horn and G. Ward, 515−537. Malden, Mass.: Blackwell Publishing Ltd.

Jackendoff, Ray. 1972. *Semantics in Generative Grammar*. Cambridge, Mass.: MIT Press.

Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59: 434−446.

Jasinskaja, Ekaterina. 2010. Modelling Discourse Relations by Topics and Implicatures: The Elaboration Default. In *Constraints in Discourse 2*, ed. P. Kühnlein, A. Benz, and C. Sidner, 61−79. Amsterdam: John Benjamins.

Ladd, D. Robert. 1980. *The structure of intonational meaning: evidence from English*. Bloomington: Indiana University Press.

Ladd, D. Robert. 2008. *Intonational phonology*. Cambridge, UK: Cambridge University Press.

Nilsenová, Marie. 2006. *Rises and Falls: Studies in the Semantics and Pragmatics of Intonation*. Doctoral dissertation, University of Amsterdam.

Pierrehumbert, Janet, and Julia Hirschberg. 1990. The Meaning of Intonation in the Interpretation of Discourse. In *Intentions in Communication*, ed. P. Cohen, J. Morgan and M. Pollack, 271−311. Cambridge MA: MIT Press.

Reese, Brian, Pascal Denis, Nicholas Asher, Jason Baldridge, and Julie Hunter. 2007. Reference Manual for the Analysis and Annotation of Rhetorical Structure.

Roberts, Craige. 1996. Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. *OSU Working Papers in Linguistics* 49: 91−136.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1: 75−116.

Schubiger, Maria. 1958. *English Intonation: Its Form and Function*. Tübingen: Max Niemeyer Verlag.

Silverman, Kim E. A., Mary Beckman, Janet Pierrehumbert, Mari Ostendorf, Colin Wightman, and Patti Price. 1992. *ToBI: A standard scheme for labeling prosody*. Paper presented at the Proceedings of the 2nd International Conference on Spoken Language Processing, Banff, Canada.

Tyler, Joseph. 2012. *Discourse Prosody in Production and Perception*. Doctoral dissertation, University of Michigan, Ann Arbor.

Department of Literature and Linguistics
Qatar University
P.O. Box 2713
Doha, Qatar
josephctyler@gmail.com