



Non-question rises in narratives produced by mothers and daughters

Meghan E. Armstrong¹, Page Piccinini², Amanda Ritchart²

¹University of Massachusetts Amherst (USA)

²UC San Diego (USA)

armstrong@umass.edu, piccinini@ucsd.edu, aritchart@ucsd.edu

Abstract

In recent years, a great deal of attention has been paid to the use of rises on non-question utterances in American English, as well as other English varieties. However, little attention has been paid to this phenomenon in the speech of children acquiring American English (AmEng). Here we analyze elicited spontaneous speech from four mother-daughter pairs in Central Connecticut. Both mothers and daughters were asked to tell a short story to the experimenter, since the “uptalk” phenomenon has been shown to be common in narratives. Results suggest that input plays a role when choosing between tunes known to convey non-finality in AmEng: mothers using more rises than level tones have daughters who do the same, and vice versa. Such an effect is not found when comparing rises to falls: daughters in general produce more rises than falls when compared to their mothers. Finally, daughters produce rises that are longer and shallower than their mothers’. Results are compared to previous work on non-question rises in AmEng. In addition, we consider implications for language development as well as language change.

Index Terms: rises, uptalk, prosody, intonation, child speech, narrative speech, slope, F0

1. Introduction

The use of rising intonation in non-question utterances, particularly when the speaker is asserting information to the hearer (rather than requesting information from the hearer) has been of great interest to both scholars and the popular media. This is perhaps because the same tunes used for such utterances (e.g., L*H-H%, H*H-H%) may be strongly associated with polar, or yes-no, questions in American English (AmEng) [1], causing the use of a rise when the speaker is asserting information to be seen as somewhat controversial. However, as [2] notes for high rising terminals (HRTs, a similar phenomenon in New Zealand English), the hearer can make use of context to infer when the rise indicates a question versus when it does not. This phenomenon, popularly referred to as *uptalk* in AmEng, has stereotypically been associated with women’s speech [3, 4, 5]. However, it is possible that the difference in frequency of uptalk, or non-question rises, in the speech of AmEng-speaking men and women is now leveling out, since [6] found no effect of gender in college students from Massachusetts and Southern California.

While social factors have indeed been of interest with respect to this phenomenon, little attention has been paid to its use in child speech. Anecdotally, while this speech pattern is often associated with “Valley Girls” [4], it is also

referred to as “childlike” speech, though research on children’s use of rising intonation when asserting information is limited. Work by [2] investigated the use of rising intonation on utterances that assert (rather than question) information in the speech of New Zealand 4- and 9-year-olds. [2] found that HRTs were a prominent feature of the 9-year-olds in her study, and to a lesser extent in the speech of 4-year-olds. The pattern was most favored in narrative contexts, which had also been shown to be the case for HRT production in the speech of Australian adults [7]. For the case of HRTs in Australian English, [7] suggest that this intonational strategy is a way to check for listener comprehension. [8] confirm that listeners are to some extent metalinguistically aware of the use of polar question intonation on non-questions as a way of checking for listener comprehension in AmEng and Majorcan Catalan. Interestingly, Majorcan Catalan uses a falling tonal configuration for polar questions that can also be exploited on non-questions in contexts such as narratives or giving directions. On the other hand, [9] observes that almost every utterance in the narratives of AmEng-speaking preadolescent boys showed rises, while the introductions and conclusions never did, suggesting that boys this age associate rises with a basic “non-finality” meaning, and falls with a “conclusion” meaning.

In any case, children must become aware of the fact that rises in AmEng are used both for seeking information (polar questions) and asserting information (non-question rises). Observing their production and comparing it to the speech of parents could tell us something about this awareness. We also expect that to a certain degree, children’s knowledge about the use of rises depends on the input they are exposed to, namely the speech of their parents. Children’s production of linguistic variables in variation has been shown to reflect that of their parents by the age of four. For instance, [10] showed that before the age of four, both of the Puerto Rican Spanish-acquiring children in her study exhibited the same patterns of variation as their parents for final boundary truncation. Thus, rise use in parents could predict rise use in children.

Here we examine the use of rises on non-questions produced in narratives by girls aged 6-7 and their mothers. Prior work has shown that the input provided by mothers to female children versus male children can differ [11], so we thus limit this study to the speech of female parents and their daughters. With this study, we aim to (1) determine the frequency distribution for rises vs. level or falling tones in the speech of mothers versus daughters and (2) determine how similar mothers’ versus daughters’ rises were produced in terms of their phonetic implementations.

2. Methods

2.1. Speakers

Four mother-daughter dyads were recorded (8 speakers total). All participants were monolingual speakers of the variety of AmEng spoken in Central Connecticut (Hartford County). Three of the daughters were six years old, and one was seven (Daughter 1). Mothers were all in their mid to late thirties, and were born and raised in the Hartford area. All participants received compensation for their participation.

2.2. Task and Procedures

All recordings took place in the families' homes, using a Zoom H4N recorder. Each participant wore a Shure WL183 condenser microphone, and recordings were done at a 48 kHz sampling rate with 16-bit quantization. While this paper focuses on an elicited narrative task, other tasks were recorded as part of a larger project on intonation in mother-child dyads.

In order to elicit narratives, the experimenter, a female from the same dialect area as the participants, asked participants to tell her a story. In all instances, daughters volunteered to tell a story first, followed by their mothers. Participants were instructed to tell the experimenter something that had happened to them, such as their last birthday party or vacation, or what happened in their favorite movie. Mothers tended to talk about vacations, and daughters tended to talk about birthday parties and movies. Since the stories were not considered new information for mothers talking to daughters (or vice versa) the story was directed to the experimenter in all cases. Thus, we do not classify the narratives as child-directed speech. While the speech was in fact adult-directed, we do believe that "child present" adult-directed speech may differ from adult-directed speech when no children are present; therefore we refer to the type of speech produced by mothers in these narratives as adult-directed, child-present speech. In some cases, speakers told two short stories in order to have an average recording time of three minutes.

2.3. Annotations and Measurements

Annotations and measurements largely follow the methodology from [6]. IPs (intonational phrase boundaries with roughly break index 4 according to MAE_ToBI labeling conventions [12]) were annotated by one transcriber and verified by a second transcriber. For each IP, the utterance contour shape (rising, level, or falling), utterance discourse type (non-question, question, or filler), and the "story mode" of the speaker (in or out of story) were coded. Contour shapes were labeled according to the following criteria: "rise" was defined as rising F0 between the nuclear stressed syllable and the end of the utterance, "level" was defined as sustained pitch, or no more than 10 Hz difference between these two points, and "fall" was defined as falling F0 between these two points. Both pitch tracking in Praat [13] and auditory perception were used to assign contour shape categories. Discourse types were decided according to the following criteria: a question was defined as an utterance that could plausibly be answered with a "yes" or "no", or if syntactically marked with a relevant *wh*-word; a filler was defined as the use of strictly filler words (e.g., *um*, *uh*), a conjunction (e.g., *and*, *but*), or the combination of the two; all other utterances were considered non-questions. Story mode of the utterance was defined according to the following criteria: out-of-story mode was defined as any instance in which the speaker was

not telling the narrative, e.g., if the speaker was asking a question, being interrupted, or answering brief follow-up questions.

Additionally, for each IP-final rise, several measurements were taken: (1) duration of rise from onset of rise until its peak, (2) F0 at onset of rise (F0 minimum), and (3) F0 at the peak of the rise (F0 maximum). All F0 measurements were transformed to ERB ($ERB = 21.4 * \log_{10}(Hz * 0.00437 + 1)$) to account for mother-daughter vocal tract differences [14]. Based on these measurements, two additional variables were computed: (1) pitch excursion (difference between the F0 maximum and minimum in ERB), and (2) slope of the rise (pitch excursion divided by the duration of the rise from minimum F0 to rise peak). Examples of the labeled rises are given in Figures 1 and 2. The onset of the rise and the rise peak are marked as *onset* and *peak*, respectively.

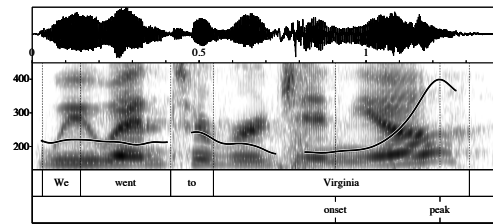


Figure 1: *Non-question rise produced by Mother 2*

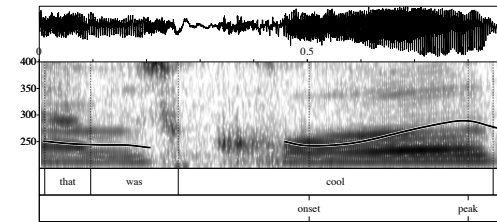


Figure 2: *Non-question rise produced by Daughter 3*

2.4. Statistical Analyses

To test for significant effects, logistic and linear mixed effects models (LMEMs) were run. Due to data scarcity, questions and fillers were not included in any of the models (see Table 1 for a summary of number of tokens), nor were utterances produced in out-of-story mode, since they were not considered part of the narrative. The logistic regressions were used to test whether there was a difference in frequency of rises based on group (mother versus daughter). Two models were run, one with the dependent variable comparing usage of final rises versus final level pitch, and one with the dependent variable comparing usage of final rises versus final falls. For both models there was only one fixed effect, group (mother, daughter). Speaker was included as a random intercept nested in dyad. This was the maximal, uncorrelated random-effects structure that converged.

Linear regressions were used to test whether there were differences in the duration and slope of rises based on group. Two models were run, one with the dependent variable as rise duration in milliseconds, and one with the dependent variable as rise slope in ERB per millisecond. For both models there was only one fixed effect, group (mother, daughter).

Speaker was included as a random intercept nested in dyad. This was the maximal, uncorrelated random-effects structure that converged. For all models, significance of fixed effects was assessed using model comparison. Alpha was set at $p < 0.05$.

3. Results

3.1. Frequency of Rises

A summary of the counts for number of tokens by utterance type and group for types of final tones is shown in Table 1.

Table 1. *Number of tokens per each variable. Percentages per group are in parentheses.*

Utterance Type	Rise	Level	Fall	TOTAL
Daughters	Rise	Level	Fall	TOTAL
Non-question	87 (35%)	106 (43%)	54 (22%)	247 (100%)
Filler	5 (9%)	42 (78%)	7 (13%)	54 (100%)
Question	2 (100%)	0 (0%)	0 (0%)	2 (100%)
Mothers	Rise	Level	Fall	TOTAL
Non-question	60 (24%)	93 (37%)	101 (40%)	254 (100%)
Filler	1 (3%)	24 (77%)	6 (19%)	31 (100%)
Question	1 (33%)	0 (0%)	2 (67%)	3 (100%)

For the model examining use of rising versus level IP-final pitch movements, there was no significant difference in the use of rising versus level tones and no significant effect of group. For the model examining the use of rising versus falling IP-final pitch movements, there was no significant difference in the use of rising versus falling tones. However, there was a significant effect of group, with daughters using significantly more rises than falls than their mothers [$\beta = 1.00$, $SE = 0.25$, $\chi^2(1) = 7.99$, $p < 0.01$]. A summary of these results are presented in Figure 3. It is interesting to note that while significant effects were only found for rising versus falling tones, the rising versus level tones appear to show more of an effect of dyad. Mothers who produce more rises (compared to level tones) have daughters who produce more rises, and mothers who produce more level tones have daughters who produce more level tones. The same effect is not seen for rising versus falling tones.

3.2. Duration and slope of rises

For the model examining rise duration, there was a significant effect of group, with daughters having longer rises than their mothers [$\beta = 153.80$, $SE = 40.00$, $\chi^2(1) = 11.09$, $p < 0.001$]. For the model examining rise slope, there was a significant effect of group, with daughters having a less steep rise than their mothers [$\beta = -0.004$, $SE = 0.001$, $\chi^2(1) = 14.84$, $p < 0.001$]. See Figure 4 for a summary of the results.

Additionally, we tested how correlated rise duration and rise slope were in order to test whether our two quantitative dependent variables were truly measuring different aspects of rises. A simple linear regression found the two variables to be highly negatively correlated, with a longer duration resulting in a shallower slope [$r = -0.40$, $p < 0.001$].

Closer inspection of the data reveals that while both groups have a negative correlation, the slope of the correlation is much less steep for daughters. See Figure 5 for a summary of the result.

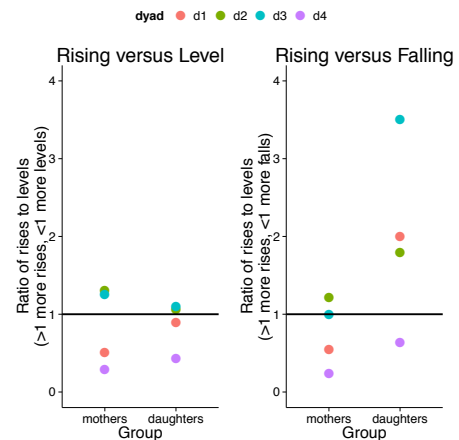


Figure 3: *Ratio of rises to level and falling tones separated by group and dyad.*

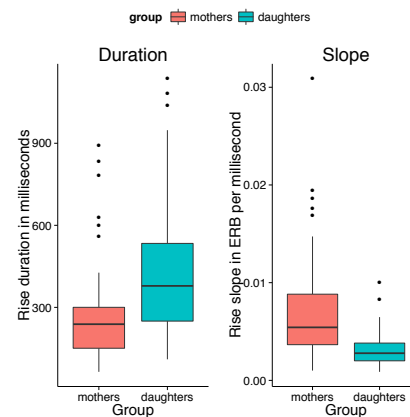


Figure 4: *Rise duration (ms) and slope (ERB per ms) separated by group.*

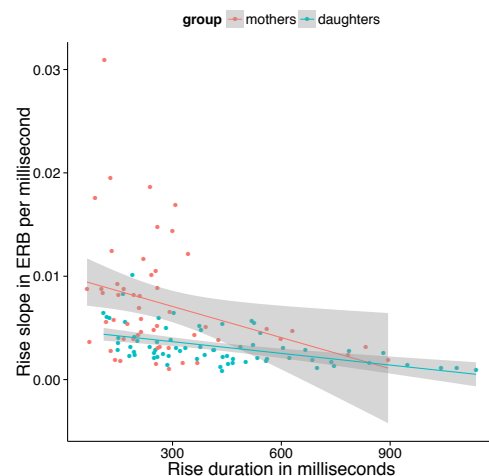


Figure 5: *Correlation of rise duration and rise slope by group.*

4. General Discussion

A variety of similarities and differences were observed in the mother-daughter dyads analyzed in this study. Our data show that while daughters do use significantly more rising IP-final pitch movements than their mothers (when compared to falling pitch), it is not the case that they are exclusively using rises in non-final position, as [9] showed. In fact, even though the daughters used significantly more rises than falls than their mothers did, both mothers and daughters are similar to [6]'s findings in that the most favored contour is a level one. For mothers, the order of frequency is Level>Falling>Rising while for daughters it is Level>Rising>Falling. This is interesting in light of [6]'s findings, since they show that college students from the same region use rises and falls at nearly the same frequency in narratives. If these children are participating in a change in progress, this would be the expected pattern.

We also find that mothers who produce more rises compared to level tones have daughters who produce more rises, and mothers who produce more level tones (compared to rising tones) have daughters who produce more level tones. The same effect is not seen for rising versus falling tones. In general, daughters always produced more rises versus falls when compared to their mothers. We find a tendency for mothers' ratios to be <1 (more falls) and daughters to be >1 (more rises). The exception to this rule was Daughter 4, whose mother produced the smallest amount of rises in the corpus. The asymmetry we find for the effect of dyad when considering rises compared to level tones versus the lack of such an effect when considering rises versus falls is quite interesting since both rises and level tones have been referred to as conveying *non-finality* in AmEng [15, 16]. Thus, we find daughters behaving more like their mothers when there is a choice between two types of non-finality: rising vs. level.

We would like to suggest that when comparing the rise versus the level tone, that the rise is the more informative of the two tones, since it can be used by the speaker to confirm that the hearer is following or understanding the speaker. We might compare this function of the rise to a discourse marker like *you know?* which has been claimed to convey a "sense of speaker involvement" [17, pp. 69]. If rises are indeed a more informative alternative to level tones at non-concluding points of a narrative (i.e., rises convey non-finality in addition to checking for comprehension), then we might say that mothers who tend to opt for the more informative option have daughters that tend to do so as well.

When considering rising versus falling tones, this is not a choice between two ways of conveying non-finality, since falls are not known to be associated with non-finality in AmEng. However, falls were used in narratives before the speaker had concluded, with mothers using more falling tones mid-narrative than their daughters in general. Daughters still demonstrated their sensitivity to non-finality and finality, though, since all daughters used falls for the concluding utterance of their narratives. While we did not analyze the pragmatic functions of non-concluding falls in the data, we propose that this difference is due to a more rudimentary knowledge of tone distribution for the daughters: they reserve fall use more for concluding rather than exploiting them for other, perhaps more sophisticated, pragmatic reasons throughout the narrative. A closer look at the pragmatic functions of non-concluding falls in narratives could be informative.

In terms of the phonetic implementation of rises, we find differences between mothers and daughters as well. In

general we can say that daughters produced longer, shallower rises than their mothers. However, even when daughters produced short rises, they were quite shallow as well, as can be seen in Figure 2. There was a negative correlation between duration and slope, so that in general longer rises tended to be shallower in general, as also found in [6]. The slope of the correlation, however, was much less steep for daughters, indicating that daughters tend to produce shallow rises even when they are short.

An initial hypothesis for the lack of steep slopes by daughters might be that they lack the articulatory control to produce steep rises, since steep rises are also faster rises. We did not calculate speech rate, but overall, daughters produced longer IPs than their mothers did, with an average duration of 1.08 seconds for mothers versus 1.16 seconds for daughters. It could thus be the case that having longer IPs affects the slope of the daughters' rises. However, in an additional task designed to elicit polar questions (not analyzed here), it does appear that daughters are able to produce steeper rises when producing polar questions. Given this data, it would be interesting to assess whether mothers produce steeper rises in polar questions despite the use of steep non-question rises in narratives. If this were the case, then it could be that children are sensitive to this question vs. non-question distinction, such that daughters are embellishing the distinction by keeping their non-question rises quite shallow and more plateau-like, as seen in Figure 2, with sharper rises reserved for questions.

Another important point is that rises in AmEng seem to be exploited quite frequently by speakers and have uses that are highly context-sensitive [8]. As less-experienced language producers, daughters might be being maximally informative in their phonetic realization of non-question rises by using a more shallow slope, thus providing more cues to the hearer that they are not to be interpreted as questions, while mothers allow the hearer to rely more on context, which could perhaps be considered a more economical strategy. In any case, it will be crucial to compare the phonetics of question rises versus non-question rises in the speech of both mothers and daughters, and more broadly parents versus children, to fully understand the discourse licensing restrictions for the different rise types for the two groups.

Comparing these results to those from [6], an interesting picture emerges, such that non-question rises are more frequent in the young daughters' speech compared to New England college-age females. In addition, New England college-age females produce more rises than New England mothers in their 30s. This trend does seem to indicate some kind of change in progress. In her work on HRTs in the speech of New Zealand 9-year-olds, [2] pointed out that it is difficult to tease out whether the higher frequency of rises in child speech can be explained by developmental reasons or whether it is a sign of change in progress. We suggest that these two factors might not be mutually exclusive of each other. If child development favors a specific change, this could perhaps keep the change "healthy." [18] points out that "the workings of dialect acquisition and change beginning in early childhood and extending into adolescence and beyond may provide fruitful avenues of research for some time to come" (pp. 133), and we believe that language change in the domain of intonation could be particularly fruitful for this line of research.

5. Acknowledgements

This research was made possible by a University of Massachusetts Amherst Faculty Research Grant/Healey Endowment Grant. We are very grateful to the mothers and daughters who participated in this study.

The second and third authors contributed equally to this work.

6. References

- [1] J. Pierrehumbert and J. Hirschberg, "The meaning of intonational contours in discourse," In P. Cohen, J. Morgan and M. Pollack (eds.), Cambridge, MA: MIT Press, pp-271-311.
- [2] H. Ainsworth, "The emergence of the high rising terminal contour in the speech of New Zealand children," *Te Reo*, vol. 37, pp. 3-20, 1994.
- [3] D. Britain, "Linguistic change in intonation: the use of high rising terminals in New Zealand English," *Language Variation and Change*, vol. 4, pp. 77-104, 1992.
- [4] A. Ritchart and A. Arvaniti, "The form and use of uptalk in Southern Californian English," in *Speech Prosody 2014, May 20-23, Dublin, Ireland, Proceedings*, 2014, pp. 331-335.
- [5] J.C. Tyler, "Expanding and mapping the indexical field: Rising pitch, the uptalk stereotype and perceptual variation," *Journal of English Linguistics*, vol. 43, no. 4, pp. 284-310, 2015.
- [6] M.E. Armstrong, P. Piccinini and A. Ritchart, "The phonetics and distribution of non-question rises in two varieties of American English," in *ICPhS 2015 – The 18th International Congress of Phonetic Sciences, August 10-14, Glasgow, Scotland, Proceedings*, 2015, paper number 0927.
- [7] G. Guy, B. Horvath, J. Vonwiller, E. Daisley and I. Rogers, "An intonational change in Australian English," *Language in Society* vol. 15, issue 1, pp. 23-51, 1986.
- [8] M.E. Armstrong & M.M. Vanrell, "Inferential constraints on question intonation in American English and Majorcan Catalan," submitted, *Speech Prosody 2016, May 31-June 3, Boston, USA, 2016*.
- [9] A. Kortenhoven, "Rising intonation in children's narratives," ms, Stanford University.
- [10] M.E. Armstrong, "The development of yes-no question intonation in Puerto Rican Spanish," Columbus, Ohio, Ohio State University dissertation, 2012.
- [11] P. Foulkes, G. Docherty, D. Watt, "Tracking the emergence of structured variation – realisations of (t) by Newcastle children," *Leeds Working Papers in Linguistics*, vol. 7, pp. 1-23.
- [12] M.E. Beckman, J. Hirschberg and S. Shattuck-Hufnagel, "The original ToBI system and the evolution of the ToBI framework," In *Prosodic Typology: The Phonology of Intonation and Phrasing*. S-A. Jun (ed.), Oxford: Oxford University Press, pp. 9-54, 2005.
- [13] P. Boersma and D. Weenink, Praat: doing phonetics by computer [Computer program]. Version 5.4.04, retrieved 10 November 2016 from <http://www.praat.org/>.
- [14] J.E. Huber, E.T. Stathopoulos, G.M. Curione, T.A. Ash and K. Johnson, "Formants of children, women and men: The effects of vocal intensity variation," *The Journal of the Acoustical Society of America*, vol. 106, issue 3, pp. 1532-1542, 1999.
- [15] D.R. Ladd, *Intonational Phonology*. Cambridge: Cambridge University Press, 1996/2008.
- [16] D. Bolinger, "Intonation in American English," In *Intonation Systems: A Survey of Twenty Languages*. D. Hirst and A. Di Cristo (eds.), Cambridge: Cambridge University Press, pp. 45-55, 1998.
- [17] L. Torres, "Bilingual discourse markers in Puerto Rican Spanish," *Language in Society*, vol. 31, issue 1, pp. 65-83, 2002.
- [18] J. Roberts, "Going younger to do difference: The role of children in language change," *University of Pennsylvania Working Papers in Linguistics*, vol. 6, issue 2, pp. 121-136, 1999.