Gender, questions and phonetic variation in uptalk: A corpus analysis of terminal rising pitch in the Santa Barbara Corpus of Spoken American English

Joseph Tyler

Morehead State University

Author Note

Joseph Tyler is now at the Department of English, Morehead State University.

Correspondence concerning this article should be addressed to:

Joseph Tyler, Department of English, Bert Combs Building 103, Morehead State University, Morehead, KY 40351.

Telephone: (734) 834 8333

Email: josephctyler@gmail.com

Short title: Gender, questions and phonetic variation in uptalk

Abstract

This study analyzes variation in rising pitch in everyday conversation from the Santa Barbara Corpus of Spoken American English. Results show that women produce larger rises on their questions than non-questions, while men do not. Additionally, informal conversations between family and friends tended to have larger rises than task-related conversations. With questions interpreted as more social than non-questions, it appears men may inhibit or women may employ an amplification of rise spans in more socially intimate interactions. These results inform our understanding of phonetic variation in terminal rising pitch and what predicts such variation.

Acknowledgments

I would like to thank Paul Warren for feedback and sharing references, Jon Yip, Sarah Ouwayda, and Katy Carlson for feedback, Emily Holley and Dallas Cox for help coding, and two anonymous reviewers.

# Gender, questions and phonetic variation in rising pitch: A corpus analysis of terminal rising pitch in the Santa Barbara Corpus of Spoken American English

Much scholarly attention has been devoted to final rising pitch, an intonational contour also referred to by names such as High-Rising Terminal (Guy and Vonwiller 1984; Warren and Britain 2000) and uptalk (Shokeir 2008). While one function of final rises is to turn statements into yes/no questions, rises have many other meanings too: rises are used to mark non-final members of a list (Cauldwell and Hewings 1996; Ladd 1980), convey epistemic uncertainty (Nilsenová 2006), indicate speaker commitments (Gunlogson 2008), or communicate how constituents in a discourse are related (Pierrehumbert and Hirschberg 1990; Tyler 2012). And while some of this work treats rises as a single category (Gunlogson 2008; Nilsenová 2006) or as part of a rise-fall-rise contour (Wagner, et al. 2013; Constant 2012), a growing body of work looks at variation in the form and meaning of final rises. Some perception studies have examined effects of rises vs. falls (Tyler 2014) or different kinds of rises (Tomlinson Jr and Fox Tree 2011; Shokeir 2008) on comprehension, but most work has explored variation in the production of rises. Among production studies of rises, a range of dialects of English have been explored, including the English of Australia (Guy and Vonwiller 1984; Fletcher and Harrington 2001), New Zealand (Warren 2005; Daly and Warren 2001), Britain (Shobbrook and House 2003), the USA (Ritchart and Arvaniti 2013), Hong Kong (Cheng and Warren 2005) and the Falkland Islands (Sudbury 2000). Despite the range of production studies on rises, little work has been done on rises in naturally-occurring contexts. This study presents the results of an analysis of rises in the Santa Barbara Corpus of Spoken American English, a collection of natural speech produced in a range of everyday activities. Using these data allows an exploration of patterns of rise variation as predicted by factors like speaker gender or questioning. It also allows the comparison of the results with data from map tasks, the format taken by most studies of rising pitch, to see if those results generalize to conversation in more everyday situations.

Rises in production studies have been analyzed along a number of dimensions, including their frequency of production, their phonological form and their phonetic characteristics. The phonological analysis tends to code the contours as a tune composed of pitch accents and boundary tones, usually represented in an autosegmental-metrical theory with ToBI-type annotation (Silverman, et al. 1992). One finding is that Australian/New Zealand English often have high rises (Fletcher, et al. 2005), i.e. rises that start high and continue to rise, while rises in Glasgow (Mayo, et al. 1997) and Southern California (Ritchart and Arvaniti 2013) more often start low and rise to high. Phonetic measures have also been examined, generally for how they correlate with contextual factors like speaker gender, the type of utterance on which the rise is uttered (e.g. question vs. statement), age, social class and other factors. Phonetic dimensions of interest have included the size of the rise, pitch dynamism (how fast pitch changes), and alignment with phonological features.

Studies that have included gender as a predictor have found systematic differences in the rises produced by men and by women. Women have been found to produce more rises (Barry and Arvaniti 2006; Ritchart and Arvaniti 2013), to start their rises later (Warren and Daly 2000; Warren and Daly 2005; Ritchart and Arvaniti 2013) and to produce more pitch dynamism (Daly and Warren 2001). Results for speakers in Southern California, London and New Zealand showed women produced bigger rises than men (Daly and Warren 2001; Barry and Arvaniti 2006; Warren and Daly 2000; Ritchart and Arvaniti 2013). Moreover, there appears to be a gendered difference in the use of rises in reaction to situational context. Linneman (2012) analyzed the responses of contestants on the game show Jeopardy! He found that both men and women produce final rises, with women producing more overall. But most interestingly, when women are in the lead they produce more rises than their average and when men are in the lead they produce fewer rises than their average. Linneman interprets this gender-differentiated pattern to be a way that women use uptalk to mitigate their success while men use uptalk to appear competitive toward other men. The relationship between gender and uptalk is visible in both macro patterns of behavior and micro dynamics of gender in interaction.

Rises also have been shown to be produced differently on questions and statements, despite earlier claims that “there is … no phonetic difference that depends on whether the high-rising contour is used with syntactically marked questions or with statements” (Ladd 1996:121). Warren and Daly (2005) found that rises on questions started earlier than on statements. Ritchart and Arvaniti (2013) found a similar pattern in Southern California English, with speakers producing rises on questions with an earlier start, and a larger rise, than on statements. Fletcher & Harrington (2001) found that their Australian English speakers produced rises on questions that started higher and rose less compared to rises on statements. Shobbrook & House (2003) explored whether Southern British English speakers similarly start question rises higher than statement rises, and they found speakers did not. This suggests two potential differences between Australian English and British or American English: 1) Australians rise less while Americans rise more on questions compared to statements; 2) Australians start rises on questions higher than on statements while British speakers do not start at different levels. A common finding across dialects is that question rises start earlier.

There have also been interactions between gender and utterance type (question vs. statement). Fletcher & Harrington (2001) found males had larger rises for statements than questions, while women’s rises were not different in size between statements and questions. While Daly and Warren (2001) found no gender difference in the size of the rise for questions and statements, they did find that women distinguished questions from statements in pitch dynamism more than men. And Warren and Daly (2005) found an interaction between gender and utterance type on rise onset position: where men’s and women’s statement rise sizes did not differ, women’s question rises started significantly later than men’s.

Other factors found to correlate with rise shape include speaker age, social class and the rise’s position in a discourse. Warren and Daly (2005) found younger speakers make more use of rises than older speakers. Research in Australia has found that rises (HRT in their study) were more common among lower-class speakers (Guy, et al. 1986). Barry and Arvaniti (2006) compare the rises of Southern California and London speakers, finding that Californians produce more rises turn-medially while Londoners produced more rises turn-finally. And Ritchart and Arvaniti (2013) found floor-holding rises were larger than statement rises but smaller than question rises. Other discourse features with which rises have been correlated are the linguistic register or speech situation (Cheng and Warren 2005; Carmichael 2005).

While the literature above has identified a range of effects, most of the results have come from studies that used a map task (Fletcher, et al. 2002; Barry and Arvaniti 2006; Shobbrook and House 2003; Fletcher and Harrington 2001; Ritchart and Arvaniti 2013). Two studies have drawn on more naturally occurring speech through the radio (Carmichael 2005) or a range of natural speech settings (Cheng and Warren 2005), but these studies are harder to compare to the other results because both their intonation analysis and contextual factors were different. As a result, our understanding of variation in rises across a range of dialects of English is largely an understanding of rises as they emerge in a map task. It is unclear whether insights into rises in map task speech generalize to more naturally occurring speech.

As in previous studies, the rises in this study are analyzed phonetically, and differences are compared across speaker gender and questioning. In addition to the common measures of rise-start pitch, rise-end pitch, and rise-span (end minus start), rise durations were also measured and examined. Analyses examine whether gender and questioning predict rise duration, as well as how duration relates to pitch. The results discussed above, particularly from Ritchart & Arvaniti (2013) for American English, lend themselves to the following predictions: 1) question rises will be larger than statement rises, and 2) women’s rises will be larger than men’s rises. Results from other dialects show evidence of interactions between gender and utterance type, but not in consistent ways. A tentative hypothesis 3) is that there will be an interaction between gender and utterance type on rise span. Another hypothesis 4) is that question rises will start higher than statement rises, an effect found for Australian English (Fletcher and Harrington 2001) but not for Southern British English (Shobbrook and House 2003).

In sum, this paper presents a corpus analysis of terminal rising pitch in everyday conversational American English, with the goal of identifying patterns of variation in the phonetic form of rises as conditioned by contextual factors (e.g. speaker gender, questioning). These results explore systematic variation in the phonetics of terminal rises, which can help explain what rises mean by disaggregating them from a single category. Instead, the variation is treated as potentially meaningful, and could illuminate the diverse uses of rises in production and variability in their perception.

**Method**

 The spoken data used as the basis of this corpus analysis come from Disc 1 of the Santa Barbara Corpus of Spoken American English (henceforth SBC). The SBC captures a wide cross-section of everyday American English, from many regions of the United States, in everyday settings. Recorded in the early 1990s, the interactions take place in contexts like an office or a kitchen, and include environmental noise associated with activities of the speakers or from the surroundings. For example, a meeting in a law office had sounds of papers shuffling and a discussion in a kitchen had the noise of people moving and washing dishes. Some of the conversations were dyadic and some were multi-party; some were mixed-gender and some were female-only.

The SBC has two benefits that make it particularly useful for the analysis of terminal rises. First, the corpus captures a broad range of contexts and speech styles in everyday interactions, which should make it easier to generalize to American English. Second, the SBC comes with all sound files transcribed using a system that codes final rising pitch. In the transcripts, each line represents an intonation unit with terminal pitch that is either falling (marked with a period “.”), flat (marked with a comma “,”) or rising (marked with a question mark “?”). Therefore, a search for question marks returns tokens of final rising pitch as identified by the transcribers. This provides a straight-forward way of collecting a sizeable sample of terminal rises. It also provides a sample of *perceived* rises, i.e. rises determined impressionistically and independently. What is unknown from the transcription, however, is how those perceived rises are realized phonetically. There may be a wide variety of kinds of rises, varying in the amount of rise, starting and ending pitch, or “rises” that even fall in pitch.

*Acoustic measurements*

This corpus analysis focused only on those tokens that permitted an acoustic analysis. Cases where the rise was masked with environmental noise or occurred with overlapping speech were excluded. If there was some overlap earlier in the intonation unit but not during the terminal rise, i.e. the rise was still analyzable, then the token was included. Out of 949 total terminal rises on Disc 1 of the SBC, 636 were selected as measureable.

Each token was annotated to a TextGrid in Praat (Boersma and Weenink 2009) with a tier that captured the terminal rise (Figure 1). This enabled the automatic measurement of pitch using a Praat script, with confidence that the beginning and end of the coded rise corresponded to the beginning and end of the rise itself. The beginning of the terminal rise was marked at an f0 valley and the end at an f0 peak.



Figure 1: Rise annotation to TextGrid in Praat.

Concerns about segmental effects of the preceding consonant on the pitch of the rise start, so-called micro-prosody, were mitigated by allowing a small buffer following the onset consonant.

All tokens were analyzed with a Praat script to extract rise durations and pitch measurements at the start and end of the rise, using the pitch range of 100-500Hz for women and 75-300Hz for men. In order to reduce false measurements, a conservative approach was taken by setting the silence threshold higher (at 0.1, above the default of 0.03) such that noise and other non-modal voicing would be less likely to be included. Problematic tokens were assessed individually, including those returning 0Hz values from the Praat script, those noted during TextGrid annotation as requiring manual measurement, and outliers. Outliers were defined as tokens at the extremes (<75Hz, >400Hz, and with a span that was >200Hz or negative) or more than two standard deviations away from that speaker’s mean. Out of a total 636 tokens, 58 (9%) were excluded because they were unmeasurable, even with manual analysis. The remaining 578 tokens served as the basis of the corpus analysis.

 While the raw f0 measurements are on the linear Hz scale, this is problematic for an analysis of rise spans between men and women: rise spans lower in pitch would be considered smaller in Hz than perceptually equivalent spans higher in pitch. A range of scales have been used to address this, including semitones, mels, Bark and ERB. The following statistical analyses will use semitones because they are shown to perform well on pitch span matching tasks.[[1]](#footnote-1)

*Annotating the structural features*

 In addition to the phonetics of the rises, structural aspects of the rises in discourse context were coded. First, each rise was coded for utterance type, i.e. whether it functioned as a question or not. This involved listening to the rise in context and inferring whether the speaker intended the rise token as a question or not, an impressionistic coding that is similar to the coding of Fletcher & Harrington (2001:223). Questions were coded by the author and two undergraduate research assistants. The goal was to focus on the information-seeking questions that are “canonically associated with questions” (Sicoli, et al. 2014:6), excluding tags and rises that are more about checking in with an interlocutor than actually inquiring. An intercoder reliability analysis shows average pairwise agreement at 82%, with a Cohen’s Kappa score of 0.637. Tokens that at least two out of the three coders agreed were questions were treated as such, meaning that a majority agree that a rise is or is not functioning as a question.

Results

Before narrowing to the phonetically analyzed tokens (see Table 3), Table 1 presents the the distribution of all rising, flat, and falling final pitch contours on Disc 1 of the Santa Barbara Corpus.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Terminal pitch contour | Corpus (n) | Men (n), % men’s contours | % by men | Women (n), % women’s contours | % by women |
| rise | 949 | 339 (6.3%) | 36% | 610 (6.5%) | 64% |
| flat | 7284 | 2654 (49.5%) | 36% | 4626 (49.7%) | 64% |
| fall | 6456 | 2370 (44.2%) | 37% | 4078 (43.8%) | 63% |
| total | 14689 | 5363 | 37% | 9314 | 63% |

Table : Distribution of rising, flat and falling terminal pitch contours in Disc 1 of the Santa Barbara Corpus of Spoken American English used in the analysis, broken out by gender.

In this data set, women produced almost two-thirds of the speech. Proportionally, women and men produced rising, flat, and falling contours at almost identical frequencies. For both men and women, rising pitch constitutes just over 6%, flat pitch constitutes almost 50%, and falling pitch constitutes around 44% of the contours.

Within this broader context, the analysis here focused on the 578 phonetically analyzable rises. Table 2 shows speaker-specific frequencies of rises, with speaker age and gender information. The rises are not normally distributed across speakers, with some speakers contributing just one token to the analysis and others contributing 50 or more. In all, 376 tokens were from twenty-five female speakers and 202 were from sixteen male speakers.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker | Gender | Age | # rises |  | Speaker | Gender | Age | # rises |
| 1 | f | 19 | 64 |  | 30 | m | 51 | 51 |
| 17 | f | 34 | 38 |  | 25 | m | 19 | 46 |
| 19 | f | 27 | 32 |  | 15 | m | 33 | 19 |
| 16 | f | 38 | 27 |  | 6 | m |  | 14 |
| 18 | f | 28 | 23 |  | 10 | m | 36 | 14 |
| 8 | f | 33 | 21 |  | 9 | m | 34 | 13 |
| 20 | f | 31 | 19 |  | 38 | m | 47 | 11 |
| 33 | f | 25 | 18 |  | 37 | m | 26 | 8 |
| 4 | f | 30 | 17 |  | 7 | m |  | 7 |
| 21 | f |  | 15 |  | 26 | m | 30 | 6 |
| 2 | f | 30 | 14 |  | 36 | m | 51 | 4 |
| 24 | f | 17 | 12 |  | 39 | m | 41 | 4 |
| 34 | f | 50 | 12 |  | 41 | m | 45 | 2 |
| 35 | f | 26 | 11 |  | 5 | m | 36 | 1 |
| 13 | f | 24 | 10 |  | 14 | m | 23 | 1 |
| 27 | f | 90 | 10 |  | 40 | m | 70 | 1 |
| 11 | f | 31 | 8 |  |  |  |  |  |
| 28 | f | 83 | 8 |  |  |  |  |  |
| 12 | f | 19 | 6 |  |  |  |  |  |
| 29 | f | 72 | 4 |  |  |  |  |  |
| 22 | f | 21 | 2 |  |  |  |  |  |
| 32 | f |  | 2 |  |  |  |  |  |
| 3 | f | 50 | 1 |  |  |  |  |  |
| 23 | f |  | 1 |  |  |  |  |  |
| 31 | f | 23 | 1 |  |  |  |  |  |

Table : Number of tokens from each speaker in the corpus. Empty cells indicate age information was not provided with the corpus.

Of the tokens that were analyzed phonetically, 56% were coded as functioning as questions (64% of men’s tokens, 51% of women’s tokens).

Basic descriptive statistics for the pitch measurements of the rises in both untransformed Hz and semitones are listed in Table 3, broken out by gender.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | N | Min. | Mean | Max. | Standard Deviation |
| Male | Rise start (Hz; semitone) | 202 | 75; 75 | 134; 84 | 262; 96 | 29; 3.43 |
| Rise end (Hz; semitone) | 202 | 107; 81 | 170; 88 | 413; 104 | 49; 4.40 |
| Rise span (Hz; semitone) | 202 | -24; -3.50 | 37; 3.97 | 159; 17.15 | 33; 3.12 |
| Rise duration (ms) | 202 | 44 | 275 | 1016 | 178 |
| Female | Rise start (Hz; semitone) | 376 | 95; 79 | 202; 92 | 362; 102 | 41; 3.49 |
| Rise end (Hz; semitone) | 376 | 139; 85 | 260; 96 | 743; 114 | 80; 4.72 |
| Rise span (Hz; semitone) | 376 | -12; -1.39 | 58; 4.04 | 423; 16.65 | 61; 3.33 |
| Rise duration (ms) | 376 | 29 | 311 | 1679 | 229 |

Table : Descriptive statistics for phonetic features of the rises

The results in Table 3 show a wider pitch span for women than men, which is also evident on the semitone scale (mean span of 3.97 for men vs. 4.04 for women). Statistical models below will test whether gender, questioning and age are significant predictors of the pitch and duration measures.

Table 4 shows the results of a correlation analysis of the phonetic measurements in semitones.

|  |  |  |  |
| --- | --- | --- | --- |
| Pearson Correlations | Rise-end | Rise-span | Rise-duration |
| Rise-start | .824\*\* | -.037 | .087\* |
| Rise-end |  | .535\*\* | .157\*\* |
| Rise-span |  |  | .148\*\* |

Table : Correlation analysis of the pitch and duration measures. \*=p<.05, \*\*=p<.01.

The highest correlations are between rise-start and rise-end, and rise-span and rise-end. For these data, the size of a rise appears related to how high a speaker ends a rise and unrelated to how high they start. A rise’s duration is less, but still significantly, correlated with rise starts, ends and spans.

*Gender and Questioning*

Statistical models testing the predictive power of gender, questioning, and age were run using using the lmer function from the lme4 package (Bates, et al. 2013) in the software package R (R Development Core Team 2013). The statistical tests mentioned below include random effects for speaker and conversation, hierarchically structured with speakers embedded in conversations. These random effects control for inter-speaker variation and differences due to the conversation itself (perhaps due to the topic of conversation or the speakers’ level of excitement). Random slopes for gender and questioning were added to “keep it maximal” (Barr, et al. 2013). Unfortunately, the random slopes made the models fail to converge and so were removed. Results for models predicting rise-start, rise-end, rise-span and rise duration are presented in Table 5. Main effects of gender, questioning and age were entered in a model with no interaction term. The interaction effects were determined in separate models after adding the interaction term. Instead of using p-values for the mixed model results, effects were deemed significant when t > 2 (Baayen, et al. 2008).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Gender | Questioning | Gender\*Questioning | Age |
| Rise-start (ST) | ß=-7.65t=-10.3 | ß=1.38t=5.66 | ß=-1.27t=-2.48 | ß=-.07t=-3.00 |
| Rise-end (ST) | ß=-8.18t=-9.34 | ß=1.99 t=5.30 | ß=-2.66t=-3.39 | n.s. |
| Rise-span (ST) | n.s. | ß=.65 t=2.28 | ß=-1.63t=-2.73 | n.s. |
| Rise duration (log) | n.s. | n.s. | n.s. | ß=-1.54t=-2.59 |

Table : Results for statistical models with random intercepts for speaker and conversation and fixed factors listed above; pitch measures on the semitone scale, duration in log(milliseconds). Non-significant (n.s.) effects determined as t<2.

The main effects of gender on rise-start and rise-end are not particularly interesting, simply showing that women’s rises tend to start and end higher. The main effects of questioning show that questioning rises start and end higher, and rise more. This result for larger rise spans on questions is consistent with results for Southern California English (Ritchart and Arvaniti 2013). The main effects of age show that older speakers tend to start their rises lower and produce rises of shorter duration, though there is no evidence age predicts rise-span.

 One concern is that the effects in Table 5 may be conditioned by tag-questions. Tag questions were not coded as questions, which in this study were the more canonical, information-seeking questions. Each token was separately coded for whether it ended with a tag question or not. When a binary variable with tag question coding was added to the models, none of the effects changed significance. Results for tag questions showed that tags had lower rise starts (ß=1.09, t=3.04), were shorter in duration (ß=99, t=3.68), but were not significantly different in rise ends or rise spans.

The results of particular interest here, however, are the interaction effects between gender and questioning. When controlling for age effects, it appears women produce rises differently on questions vs. non-questions while men do not. This difference can be seen in Figure 2, which compares rise spans on men’s and women’s productions of questions and non-questions.



Figure 2: Bar graph of rise-span for questions and statements, clustered by gender.

Figure 2 shows males are not changing their rise spans between questions and non-questions, while women rise more for questions. To test this claim directly, the models without interactions but with questioning and age were re-run with only the data for men and then with data only for women. When restricted to men, questioning did not predict rise-start (ß=2.05, t=.89), rise-end (ß=.18, t=.31), or rise-span (ß=-.31, t=-.69), while for women questioning does (rise-start: ß=1.80, t=5.99; rise-end: ß=2.90, t=6.11; rise-span: ß=1.20, t=3.35). Women’s questioning rises start higher, end higher and rise more than their non-questioning rises, while men make no such distinction.

 Thus far, analyses have not taken into account the content of the conversations. While a micro-analysis of each rise in its interactional context is beyond the scope of this study, the descriptions of each conversation provide some insight. More specifically, it appears there is variation between conversations in terms of formality, with some between intimates (friends, family) and others being more business or task related. In Table 6, the conversations are sorted from lowest to highest mean rise span. The contextual descriptions suggest that the more formal, task-related interactions tend to have smaller mean rise span, while the more informal conversations between family and friends have larger rise spans. Also notable is that this separation cuts across genders. While conversation 1 may appear to be an exception, the presence of an outsider, a “visitor and near stranger”[[2]](#footnote-2), may have the effect of shifting the orientation of the conversation towards more formality. In this conversation, Lynne is telling Lenore about her studies. Even though it occurs in the home, the function is more task-related. So what might otherwise be a family context between intimates (with larger rises) shifts to a formal business-oriented context (with smaller rises).

A second example suggests that the size of rises may not reflect categories associated with conversational co-participants (e.g. gender, family, partners) or the location (home vs. office), but instead depend on the function of the interaction. Conversation 9 involves a girl Kathy helping her boyfriend Nathan prepare for a math test. On the one hand, they are a couple (like the pair in conversation 5), which given a purely relationship-based analysis may predict less formality. But in the conversation they are doing work, a factor that may be more important for their rises than category membership. That is, their activity focused on preparing for a math test may be more important than their relationship status as boyfriend and girlfriend. Their rise spans seem to reflect this work focus, patterning more with other more formal, task-related conversations.

Table : Descriptions of each of the fourteen conversations analyzed, sorted from smallest to largest mean rise-span.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Conversation | Gender |  | Mean rise-span(combined) | Mean rise-span(m) | Mean rise-span(f) | #Quot | #Rises | Task vs Social | Context[[3]](#footnote-3) |
| 10 | 2m,1f | 3 | 2.16 | 2.16 |   |   | 6 | Task | business (arts society)  |
| 8 | 1m,3f | 4 | 2.81 |   | 2.81 |   | 37 | Task | task-related (legal) |
| 1 | 3f | 3 | 3.1 |   | 3.1 |   | 73 | Social | family (daughter, mother, visitor) |
| 9 | 1m,1f | 2 | 3.2 | 2.92 | 4.27 |   | 58 | Task | task-related (math), a couple |
| 7 | 2f | 2 | 3.33 |   | 3.33 | 1 | 55 | Social | sisters  |
| 14 | 4m,1f | 5 | 3.35 | 3.35 |   |   | 18 | Task | task-related (bank) |
| 4 | 2m,4f | 6 | 3.66 | 2.37 | 3.72 | 2 | 25 | Social | family (sisters) |
| 3 | 2m,1f | 3 | 3.86 | 3.44 | 4.41 | 1 | 48 | Social | family/friends (married couple and friend)  |
| 12 | 5m,4f | 9 | 3.94 | 3.88 | 4.98 |   | 54 | Task | university lecture |
| 2 | 3m, 1f | 4 | 4.07 | 4.63 | 3.36 |   | 39 | Social | family/friends (married couple + two friends)  |
| 13 | 2m,3f | 5 | 5 | 5.25 | 4.92 | 1 | 53 | Social | family (two parents, two kids and spouse) |
| 6 | 2f | 2 | 5.18 |   | 5.18 |   | 44 | Social | family (cousins, lively) |
| 11 | 3f | 3 | 5.29 |   | 5.29 | 1 | 22 | Social | friends |
| 5 | 1m,1f | 2 | 6.32 | 7.17 | 5.72 | 15 | 46 | Social | a couple |

[[4]](#footnote-4) This model returns almost significant (ß=-.95, t=-1.94)

To test whether conversational context as task-related or social predicts rise span, a mixed model was fitted to the data. The model had a binary fixed factor coding whether the conversation was task-related or social (see Table 6), covariates for age, gender and questioning, and random effects for speaker and conversation, hierarchically structured with speakers embedded in conversations. This model returned a trend for the task-related vs. social nature of the conversation predicting risespan (ß=1.08, t=1.83). When conversation 1 is recoded as task-related (as discussed above), the effect becomes significant (ß=1.06, t=2.01). While it is clearly not the only factor that matters, more informal, social conversations do seem to have larger rise spans than more task-related ones.

While the formal/informal continuum may predict rise spans, it does not exactly explain what people are doing in the conversations differently when speaking formally or informally. On the one hand, it may be a simple factor of emotional expressivity, where formal interactions require more composure than informal ones and so have reduced pitch range overall, which would then be visible in smaller rises. One factor in this corpus analysis that could speak explain the formality dimension is the distribution of quoted speech rises. Overall, quoted speech rises seem to occur more often in more informal, intimate interactions and, perhaps not coincidentally, also had the largest rise spans.

One conversation that stood out is conversation 5, which alone had a majority of the rises in quoted speech. This distribution of quoted speech rises is important because these rises sometimes go extremely high. In the clip whose pitch contour is shown in Figure 3, the speaker is telling a story about someone who bothered the speaker by jumping a lot. Then, she quotes what she herself said in the story context. The speaker is animated and excited, and reaches a peak of nearly 800Hz.



Figure 3: The largest rise in the corpus (6\_17).

The extreme pitch excursion seems related to the fact that she is animating quoted speech as part of a story.

The intonation of quoted speech has been argued to be characterized by a shift to a higher pitch register (Wennerstrom 2001), though local interactional needs may complicate this generalization (Klewitz and Couper-Kuhlen 1999). An analysis of this corpus of rises found that quoted speech tended to start and end higher, and rise more overall, relative to non-quoted rises (Tyler 2014), which suggests that a general pattern of quoted speech being produced higher in pitch is also evident in final rises. Quoted speech may show up more often in informal interaction, which could help explain why more informal conversations seem to have larger rise spans.

**Discussion**

The results presented above show variation in the form of rises from an analysis of everyday conversational data from the Santa Barbara Corpus of Spoken American English (SBC). There was a main effect of questioning on rises, with rises functioning as questions having higher rise-starts, higher rise-ends and larger rise-spans (fulfilling prediction 1). Previous results are mixed on whether speakers start their rises higher on questions than statements, with Australian English speakers doing so (Fletcher and Harrington 2001) but Southern British English speakers not doing so (Shobbrook and House 2003). The results in Table 5 show speakers in this study start their questioning rises higher than non-questioning rises (fulfilling prediction 4). So while both US and Australian English speakers start their questioning rises higher, Australian English speakers rise less while US speakers rise more. The result for questioning on rise span fits with previous findings for Southern California English (Ritchart and Arvaniti 2013), which found that question rises were twice as large as statement rises. By contrast, Southern British English (Shobbrook and House 2003) showed no such difference between questioning and non-questioning rises and Australian English actually showed the opposite effect, with questioning rises starting higher but rising *less* compared to rises on statements (Fletcher and Harrington 2001). This suggests the relationship between rise span and discourse function (e.g. question vs. non-question) varies between American, British and Australian dialects of English.

Unlike previous research, which found women produce larger rises than men in American, British and New Zealand English (Barry and Arvaniti 2006; Warren and Daly 2000; Ritchart and Arvaniti 2013), this study showed no main effect of gender on rise-span (against prediction 2). There were, however, significant interactions between gender and questioning, with women producing larger rises on questions than non-questions while men make no such distinction (fulfilling prediction 3). This contrasts with the results of Ritchart and Arvaniti (2013), who found no interactions between gender and discourse function. Their study and this one differed by speech situation, a map task in a laboratory context for Ritchart and Arvaniti and everyday contexts like the living room or kitchen for this study. They also differed by how discourse function was classified: Ritchart and Arvaniti used four distinct categories (question, statement, holding the floor, confirmation request) while this study contrasted just questions and non-questions. And finally, Ritchart and Arvaniti focus on Southern California English, while the Santa Barbara Corpus has speech from around the US. It seems possible the relationship between gender, discourse function and final rises could vary by speech context, discourse function coding and region.

Overall, three of the four predictions were fulfilled and one was not (a gender effect on rise span). This suggests that map task results generalize fairly well to naturally occurring speech, with the exception of a simple main effect of women producing larger rises not showing up. A second difference between these results and map task results comes in the frequency of rise production, with men and women each producing just over 6% of their intonation units with final rising pitch. This contrasts with previous research that found women produce more rises than men (Barry and Arvaniti 2006; Ritchart and Arvaniti 2013; Linneman 2012). Given that much previous research has looked at rising pitch production with map tasks in laboratory contexts, these results with speech in more natural settings suggest frequency of rise production and gender effects on rise size may differ between the two contexts.

There are many unresolved questions related to rising pitch. For example, rises analyzed here were only those that were phonetically analyzable. Perhaps rises spoken during overlapping speech or in noisy contexts are different from those analyzed here. It is difficult to assess how the production of rises is changing over time or across dialect groups (though see Warren 2005). There are also many open questions about how differences in production are perceived, or even ways the exact same contour could be perceived differently by different listeners. Existing research has shown differences in perception based on listener gender (Shokeir 2008), age and regional origin (Tyler 2013). A fuller explanation of why there are production differences by gender and utterance type may depend on an understanding of how perception varies across these and other dimensions. And finally, this analysis focused solely on those final contours coded by the transcribers as rises. Results could potentially vary if the analysis were restricted to tokens defined by other means, e.g. f0 contours.

The results in this study and much prior work on gender and rises have treated gender as a universally relevant macro-social category. This work often employs a binary male/female gender classification and then analyzes correlates of each. More recently, research on language and gender has shifted away from treating each person as male or female to construing gender as constructed out of locally relevant stances and styles (Eckert and Podesva 2011; Cameron 2011). The construct of gender is built out of social actions, and macro correlates of gender categories may be less about the categories themselves than the fact that members of those categories are more likely to perform some social actions compared to other social actions. One way to analyze gender in such a paradigm is to give more in-depth study to stylistic variation in individual contexts. An example of a micro-analysis of rising pitch comes from Podesva (2011), who presents a case study of how a gay professional exploits final rises differently between professional and social contexts: the speaker produces more rises in professional settings compared to social settings, though the acoustics of his rises do not differ between settings. Ideally, micro and macro approaches to the meaning of rising pitch can inform each other, identifying how rising pitch is exploited in individual contexts and how it helps construct macro categories like gender and sexual orientation.

While mostly macro in nature, this study examined variation in rise spans by the type of interaction in each conversation. The conversations varied from more task-related, business-oriented interactions on the one hand, to more informal, intimate or familial interactions on the other. In general, larger rise spans tended to be in conversations in the more informal, intimate conversations, while task-related conversations had smaller rises. This pattern is reinforced by comparing the two conversations between intimates, which in terms of social categories (boyfriend, girlfriend) are the same. In conversation 9, a woman was helping her boyfriend prepare for a math test; in conversation 5, a couple was laying in bed chatting. Conversation 9 had relatively small rises, while conversation 5 had the largest rises of all conversations, suggesting that the task-related nature of the former and the intimate nature of the latter led to smaller and larger rises respectively. It seems that the nature of the interaction (studying vs. chatting casually) matters more for the size of the rises than interlocutors’ relationship status. This analysis supports the claims of Ochs (1992) that linguistic correlates of social categories (e.g gender, relationship status) may be more about differences in the frequency with which members of each gender perform actions that correlate with those linguistic features than anything inherent to the categories themselves. This suggests that other studies’ findings that women produce larger rises could be more about women tending to interpret interactions in more social terms. Moreover, the result that women had larger rises on questions than non-questions while men did not may be about the sociality of questioning. By creating an expectation for a response, questions involve an interlocutor in ways statements do not. The gender difference in rise spans on questions vs. non-questions may be explained as an inhibition by men to fully express the social nature of questions, or an embrace of the sociality of questioning by women.

One of the most important characteristics of terminal rising pitch in the USA is the stigma is carries. Popular media have been particularly interested in rises, i.e. “uptalk”, sometimes serving as a forum for criticizing the form or its users (Hogenboom 2013; Gorman 1993; Davis 2010; Horowitz 2006; Quenqua 2012). A more thorough understanding of variation in the production and perception of uptalk could help us better understand, and potentially mitigate, the stigma associated with the form. Though people tend not to discuss detailed variation in the form of a rise when criticizing uptalk, the production data discussed here show that women are doing more with their rises than men: in these data with American English, women’s rises are different in questions vs. non-questions while men’s are not. A pattern of treating rises as a single category misses this meaningful variation, and labeling rising pitch in general as “question intonation” erases (Irvine and Gal 2000) both non-questioning uses of rises and variation in how questioning rises are produced.

If the expectation is that rises are to be used solely for questioning, then alternate uses will be interpreted as infelicitous or degraded instead of as simply different. We know that rises are used in socially much more complex ways than simply questioning, as McLemore (1991) showed in her ethnographic study of young women’s speech in a sorority. The result of the mismatch between ideologies of rising pitch and actual production may be miscommunication, with some speakers and listeners producing and perceiving rises in different, and potentially more complex, ways than others.

 In conclusion, this corpus study has shown systematic variation in the phonetic form of rises in everyday American English. Women, but not men, produced their question rises larger than their non-question rises, and more informal, intimate interactions had larger rises than more task-related interactions. This systematic variation suggests it is problematic to discuss rises as an undifferentiated category. In addition, finding these patterns in everyday conversational American English makes the results more generalizable beyond the laboratory contexts that have served as the basis for much previous work. Rises vary in systematic ways, and understanding that variation can help us understand the complex linguistic and social meanings they carry.

**References**

Baayen, R. H., D. J. Davidson, and D. M. Bates. "Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items." *Journal of Memory and Language* 59, no. 4 (2008): 390-412.

Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. "Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal." *Journal of Memory and Language* 68, no. 3 (2013): 255–78.

Barry, Angela S., and Amalia Arvaniti. ""Uptalk" in Southern Californian and London English." In *BAAP 2006 Colloquium*. Queen Margaret University College, Edinburgh, 2006.

Bates, Douglas, Martin Maechler, and Ben Bolker. "Lme4: Linear Mixed-Effects Models Using S4 Classes." R package version 0.999999-2. (2013).

Praat: Doing Phonetics by Computer.

Cameron, D. "Sociophonetics and Sexuality: Discussion." *American Speech* 86, no. 1 (2011): 98-103.

Carmichael, Lesley Marie. "Situation-Based Intonation Pattern Distribution in a Corpus of American English." PhD, University of Washington, 2005.

Cauldwell, Richard, and Martin Hewings. "Intonation Rules in Elt Textbooks." *ELT Journal* 50, no. 4 (1996): 327-34.

Cheng, Winnie , and Martin Warren. "// / Can I Help You //: The Use of Rise and Rise-Fall Tones in the Hong Kong Corpus of Spoken English." *International Journal of Corpus Linguistics* 10, no. 1 (2005): 85-107.

Constant, Noah. "English Rise-Fall-Rise: A Study in the Semantics and Pragmatics of Intonation." *Linguistics and Philosophy* 35, no. 5 (2012): 407-42.

Daly, Nicola, and Paul Warren. "Pitching It Differently in New Zealand English: Speaker Sex and Intonation Patterns." *Journal of Sociolinguistics* 5, no. 1 (2001): 85-96.

Davis, Hank. "The Uptalk Epidemic: Can You Say Something without Turning It into a Question?" *Psychology Today*, October 6 2010.

Eckert, P., and R. J. Podesva. "Sociophonetics and Sexuality: Toward a Symbiosis of Sociolinguistics and Laboratory Phonology." *American Speech* 86, no. 1 (2011): 6-13.

Fletcher, Janet, Esther Grabe, and Paul Warren. "Intonational Variation in Four Dialects of English: The High Rising Tune." In *Prosodic Typology: The Phonology of Intonation and Phrasing*, edited by Sun Jun: Oxford University Press, 2005.

Fletcher, Janet, and Jonathan Harrington. "High-Rising Terminals and Fall-Rise Tunes in Australian English." *Phonetica* 58, no. 4 (2001): 215-29.

Fletcher, Janet, Roger J. Wales, Lesley F. Stirling, and Ilana M. Mushin. "A Dialogue Act Analysis of Rises in Australian English Map Task Dialogues." In *Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody*, 299-302. Aix-en-Provence: Universite de Provence, 2002.

Gorman, James. "Like, Uptalk?" *The New York Times*, August 15 1993.

Gunlogson, Christine. "A Question of Commitment." *Belgian Journal of Linguistics* 22, no. 1 (2008): 101-36.

Guy, Gregory, Barbara Horvath, Julia Vonwiller, Elaine Daisley, and Inge Rogers. "An Intonational Change in Progress in Australian English." *Language in Society* 15, no. 1 (1986): 23-51.

Guy, Gregory R., and Julia Vonwiller. "The Meaning of an Intonation in Australian English." *Australian Journal of Linguistics* 4, no. 1 (1984): 1-17.

Hogenboom, Melissa. "More Men Speaking in Girls' 'Dialect', Study Shows." <http://www.bbc.co.uk/news/science-environment-25232387>.

Horowitz, Jason. "City Girl Squawk: It’s Like So Bad- It. Really. Sucks?" *New York Observer*, March 27 2006.

Irvine, Judith, and Susan Gal. "Language Ideology and Linguistic Differentiation." In *Regimes of Language: Ideologies, Polities, and Identities*, edited by P. V. Kroskrity, 35-84. Santa Fe: School of American Research Press, 2000.

Klewitz, Gabriele, and Elizabeth Couper-Kuhlen. "Quote-Unquote. The Role of Prosody in the Contextualization of Reported Speech Sequences." *Pragmatics* 9, no. 4 (1999): 459-85.

Ladd, D. Robert. *The Structure of Intonational Meaning: Evidence from English*. Bloomington: Indiana University Press, 1980.

Linneman, Thomas J. "Gender in Jeopardy!: Intonation Variation on a Television Game Show." *Gender & Society* 27, no. 1 (2012): 82-105.

Mayo, Catherine, Matthew Peter Aylett, and D. Robert Ladd. "Prosodic Transcription of Glasgow English: An Evaluation Study of Glatobi." In *Intonation: Theory, Models and Applications*, 1997.

McLemore, Cynthia Ann. "The Pragmatic Interpretation of English Intonation: Sorority Speech." 1991.

Nilsenová, Marie. "Rises and Falls: Studies in the Semantics and Pragmatics of Intonation." U. of Amsterdam, 2006.

Ochs, Elinor. "Indexing Gender." In *Rethinking Context: Language as an Interactive Phenomenon*, edited by Allessandro Duranti and Charles Goodwin, 335-58. Cambridge: Cambridge University Press, 1992.

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

Podesva, Robert. "Salience and the Social Meaning of Declarative Contours: Three Case Studies of Gay Professionals." *Journal of English Linguistics* 39 (2011): 233-64.

Quenqua, Douglas. "They’re, Like, Way Ahead of the Linguistic Curve." *The New York Times*, 27 February 2012.

R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ritchart, Amanda, and Amalia Arvaniti. "The Use of High Rise Terminals in Southern Californian English." In *166th Meeting of the Acoustical Society of America*, 060001: San Francisco, California, 2013.

Shobbrook, K, and Jill House. "High Rising Tones in Southern British English." Paper presented at the 15th International Congress of Phonetic Sciences, Barcelona, 2003.

Shokeir, Vanessa. "Evidence for the Stable Use of Uptalk in South Ontario English." Paper presented at the New Ways of Analyzing Variation 36, 2008.

Sicoli, M. A., T. Stivers, N. Enfield, and S. C. Levinson. "Marked Initial Pitch in Questions Signals Marked Communicative Function." *Language and Speech*  (2014).

Silverman, K. E. A. , M. Beckman, J. Pierrehumbert, M. Ostendorf, C. W. S. Wightman, and P. Price. "Tobi: A Standard Scheme for Labeling Prosody." Paper presented at the Proceedings of the 2nd International Conference on Spoken Language Processing, Banff, Canada, 1992.

Sudbury, Andrea. "Variation and Change in Falkland Island English." University of Essex, 2000.

Tomlinson Jr, John M., and Jean E. Fox Tree. "Listeners' Comprehension of Uptalk in Spontaneous Speech." *Cognition* 119, no. 1 (2011): 58-69.

Tyler, Joseph. "Discourse Prosody in Production and Perception." University of Michigan, 2012.

———. "The Many Meanings of Uptalk: An Examination of Perceptions of Rising Terminal Pitch on Declaratives." In *New Ways of Analyzing Variation 42 (NWAV)*. Pittsburgh, 2013.

———. "Prosody and the Interpretation of Hierarchically Ambiguous Discourse." *Discourse Processes* 51, no. 8 (2014): 656–87.

———. "Rising Pitch and Quoted Speech in Everyday American English." Paper presented at the Speech Prosody, Dublin, 2014.

Wagner, Michael, Elise McClay, and Lauren Mak. "Incomplete Answers and the Rise-Fall-Rise Contour." In *SemDial*. Amsterdam, 2013.

Warren, Paul. "Patterns of Late Rising in New Zealand English." *Language Variation and Change* 17, no. 2 (2005): 209-30.

Warren, Paul, and David Britain. "Intonation and Prosody in New Zealand English." 2000.

Warren, Paul, and Nicola Daly. "Characterizing New Zealand English Intonation: Broad and Narrow Analysis." In *Languages of New Zealand*, edited by Allan. Bell, Ray Harlow and Donna Starks. Wellington: Victoria University Press, 2005.

Warren, Paul, and Nicola Daly. "Sex as a Factor in Rises in New Zealand English." In *Gendered Speech in Social Context: Perspectives from Gown and Town*, edited by Janet Holmes, 99-115. Wellington: Victoria University Press, 2000.

Wennerstrom, Ann. *The Music of Everyday Speech:Prosody and Discourse Analysis*: Oxford University Press, 2001.

1. Daly & Warren (2001) argue for using the pseudo-logarithmic ERB scale for analyzing rises, a scale that was developed from research on prominence-lending pitch movements (Hermes & van Gestel 1991). Nolan (2011) uses a psycho-acoustic span-matching task to assess the relative strength of different scales, finding that semitones and ERBs performed best, with semitones marginally better. Given these results and the fact that this paper is looking at rise spans and not pitch prominences, this paper will use semitones. Semitones were calculated as follows, where x is the f0 value in Hz:

Semitone (ST): 12 \* log(x)/log(2)

Rise spans were calculated as 12 \* log(y/x)/log(2), where y is f0 at rise end and x is f0 at rise start. [↑](#footnote-ref-1)
2. http://www.linguistics.ucsb.edu/research/santa-barbara-corpus [↑](#footnote-ref-2)
3. Descriptions of each conversation are available on the website for the corpus: http://www.linguistics.ucsb.edu/research/santa-barbara-corpus [↑](#footnote-ref-3)
4. [↑](#footnote-ref-4)