

## **UC Merced**

# **Proceedings of the Annual Meeting of the Cognitive Science Society**

### **Title**

Talking it up: How the function of rising declaratives depends on prolongations and listeners' expectations

### **Permalink**

<https://escholarship.org/uc/item/3nz6b6xc>

### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

### **ISSN**

1069-7977

### **Authors**

Tree, Jr., Jean Fox  
Tomlinson, John

### **Publication Date**

2009

Peer reviewed

# Talking it up: How the function of rising declaratives depends on prolongations and listeners' expectations

John M. Tomlinson, Jr. (otomlins@ucsc.edu)

University of California - Santa Cruz  
Department of Psychology, SSII Rm. 273  
Santa Cruz, CA 95064 USA

Jean E. Fox Tree (foxtree@ucsc.edu)

University of California - Santa Cruz  
Department of Psychology, SSII Rm. 273  
Santa Cruz, CA 95064 USA

## Abstract

Listeners' comprehension of phrase final rising pitch on declarative utterances, or *uptalk*, was examined to test the hypothesis that prolongations might differentiate conflicting functions of rising pitch. In Experiment 1 we found that listeners rated prolongations as indicating more speaker uncertainty, but that rising pitch was unrelated to ratings. In Experiment 2 we found that prolongations interacted with rising pitch when listeners monitored for words in the subsequent utterance. Words preceded by prolonged uptalk were monitored faster than words preceded by non-prolonged uptalk. In Experiment 3 we found that the interaction between rising pitch and prolongations depended on listeners' beliefs about speakers' mental states. These results demonstrate that form/function mappings may need to take temporal information into account. They also demonstrate that form/function mappings are not predetermined, but can vary depending on listener expectations.

**Keywords:** Prosody; Intonational Meaning; Pragmatics; Speech Comprehension; Prolongations; Uptalk

## Introduction

In the following exchange, Stephanie makes a blanket statement about violence in video games. What is unusual about this declarative utterance is that it ends with something more indicative of interrogative utterances: rising pitch (\* = pitch accent, / = rising pitch):

Stephanie: I don't think\* that video games\* make  
you become violent/  
Marcus: Uh-huh  
Stephanie: Yeah, because people\* can separate  
reality\* from, uh, like, video games

Two reasons why Stephanie might raise her pitch at end of her stated opinion are that (1) she is indicating that she is unsure of the truth propositional content of her utterance, either whether she believes it or whether Marcus believes it, or (2) she believes what she is saying but has yet to provide further information bolstering her claim. As spelled out in the literature, these two functions can be differentiated as to whether they apply to the utterance she just stated ("I don't think that video games make you become violent"), a *backward-looking* function, or to the subsequent utterance

("yeah, because people can separate reality from uh like video games"), a *forward-looking* function (Allen, 1986; Pierrehumbert & Hirschberg, 1991; Fletcher et al. 2002; House, 2007).

Rising declarative pitch, or *uptalk* (McLemore, 1991), is sometimes considered a restricted dialectal variation. In fact, however, uptalk is common in Australian English (Allen, 1986), Southern British English (Cruttenton, 1998), and Canadian English (Shokeir, 2008). This suggests that uptalk is a productive linguistic phenomenon.

We propose that listeners disambiguate the proposed two functions of uptalk using (1) the temporal context of prolongations and (2) expectations about the speakers' knowledge states. In Experiment 1, we tested listeners' off-line interpretations of prolongations and uptalk with respect to speaker knowledgeableness. In Experiment 2, we tested the speed at which listeners can monitor for words following prolongations or uptalk. In Experiment 3, we tested how listeners' beliefs about speakers' knowledge states influenced their on-line word monitoring performance.

## Intonation in time

A long standing goal of intonational meaning research has been to make one-to-one mappings between forms and functions (Pierrehumbert & Hirschberg, 1991). In the autosegmental approach, intonational events are represented as tones that combine to form phrasal tunes which are interpreted similarly regardless of temporal realization. We argue, as do others, that intonation is inherently temporal (Clark, 2002; Kohler, 2004; Ramus & Mehler, 1999). As a consequence, how intonation is expressed temporally might affect its interpretation.

The function of uptalk has been elusive. Gunlogson (2001) argued that uptalk allows speakers to distance themselves from the truth-propositional content of their utterances, shifting the burden of affirming or disconfirming truthfulness to addressees. This account might explain the backward-looking function of uptalk, but it does not explain the forward-looking function. House (2007) proposed that uptalk has two functions: (1) questioning or testing the relevance of the speech act to the current discourse context and (2) indicating continuity, specifically that more information is pending. Both of these accounts highlight the *Track 2* (Clark, 1996) or *procedural* (Wilson & Wharton,

2006) meaning of rising pitch. What is unclear is how listeners assign one or the other function to a particular intonational token that they hear.

We propose that prolongations interact with uptalk to yield different interpretations. Prolongations indicate that a delay is in progress, with fewer and shorter delays following rather than preceding them (Clark & Fox Tree, 2002). Recognizing a brief delay can focus listeners' attention on upcoming talk (Fox Tree, 2001). If a prolongation co-occurs with uptalk, we predict that the uptalk will be interpreted with a forward-looking function. But if uptalk is produced without prolongation, we predict the interpretation will be backward-looking. Listeners' attention could be focused on the prior information for either reason identified above: either shifting the responsibility to listeners to determine an utterance's truth value, or alerting listeners that they should attest to the relevance of the utterance.

### Prosody and mental state inferences

Although much literature exists on the relationships between prosodic information and emotional states (see Murray & Arnott, 1993, for a review), less is known about the relationship between prosodic information and inferences about a speaker's knowledge level. In support of the hypothesis that uptalk provides listeners with information about speakers' knowledge states, listeners rated speakers' feelings of knowing lower when speakers' answers were marked with rising pitch and longer response latencies (Brennan & Williams, 1995).

We propose that prolongations and uptalk will influence not only off-line judgments of speaker knowledgeableness (as demonstrated in part by Brennan & Williams, 1995), but also on-line processing of information. There is some precedent for this hypothesis. Information about a speaker's ability to produce the names of words influenced on-line responses to what the speaker said (Arnold, Hudson-Kam, & Tanenhaus, 2007). If the speaker were considered a normal language user, hearing a disfluency prompted listeners to look at previously unmentioned items in an array of items. But if the speaker were thought to have object agnosia, listeners did not necessarily look at the unmentioned items.

More precisely, we propose that listeners will interpret prolongations and rising pitch differently depending on whether they believe the speakers to be knowledgeable or not about what they are saying.

### Overview of studies

Our main goal is to assess how listeners understand rising declaratives. Our main prediction is that prolongations will interact to differentiate two conflicting functions of rising pitch, backward-looking vs. forward-looking. We also seek to identify how listeners establish these relationships between intonational form and function.

In Experiment 1, we test the hypothesis that rising pitch signals knowledgeableness to the listener relative to falling pitch. Results from studies on the feeling of knowing predict that rising pitch should be interpreted as less knowledgeable than falling pitch. Feeling of knowing studies do not provide

a prediction for prolongations.

In Experiment 2, we test the effect of rising pitch and prolongations on the monitoring of upcoming information. When listeners expect a brief delay after hearing an *uh*, attention is focused on upcoming talk and word monitoring is faster (Fox Tree, 2001). Prolongations mark ongoing delays, anticipating briefer delays after them than before them (Clark & Fox Tree, 2002). Similar to *uhs* (Fox Tree, 2001), listeners may focus on upcoming words after prolongations; that is, prolongations may direct listeners' attention forward rather than backward. Earlier studies make no prediction for the effect of pitch on word monitoring.

In Experiment 3, we test the role of listener expectations on on-line speech comprehension. Half the participants were told that speakers had memorized facts from cue cards and had trouble reproducing them later because they were not experts in the relative domain. Half were told that the speakers were experts in a domain related to the content of what they were saying. If listeners interpret rising pitch and prolongations with a predetermined set of rules, then reaction times should be similar between the expert and non-expert conditions. But if listeners take information about speakers into account when listening to talk, reaction times might differ. For example, if the speakers are thought to be non-experts, then rising pitch may serve a backward-looking function yielding slower reaction times. But if speakers are thought to be experts, rising pitch may serve a forward-looking function yielding faster reaction times.

## Experiment 1

Listeners rated utterance pairs on how accurate speakers were in the knowledge they were conveying.

### Methods

**Participants** Twenty UCSC students participated in this experiment in exchange for course credit. There were 12 women and 8 men. All participants were native speakers of English.

**Materials** Forty-eight utterance pairs were selected from a specially compiled corpus of spontaneous speech. Speakers created spontaneous sentence frames to convey celebrity facts to an addressee who attempted to select the celebrity out of an array. For example, upon reading *place of birth: Brooklyn*, the speaker might say to the addressee, "This actor was born in Brooklyn, New York." The utterance pairs consisted of two sentences spontaneously produced by the same speaker; for example, "I have two children. I was the Princess of Wales."

The first utterance of the pair had either rising prolonged pitch (N = 12), rising non-prolonged pitch (N = 12), falling prolonged pitch (N = 12), or falling non-prolonged pitch (N = 12). Rising and falling pitch were distinguished using ToBI guidelines (Beckman & Ayers, 1994). Rising pitch consisted of either H\* L-H% or H\* H-H%, whereas falling pitch consisted of either H\* H-L% or H\* L-L%. Prolonged and non-prolonged syllables were differentiated

qualitatively by examining the nature of the rise or fall: Falls and rises that peaked or reached their nadir at the beginning of the nucleus were considered non-prolonged, where falls and rises that didn't peak or reach its nadir until the coda were considered prolonged. Prolonged and non-prolonged syllables were also differentiated quantitatively. Prolonged syllables averaged 501 ms (SD = 156 ms; range 228 ms – 433 ms), whereas non-prolonged syllables averaged 291 ms (SD = 133 ms; range 313 ms – 940 ms).

For each utterance with a rising contour, a matched utterance was digitally created with a falling contour. Likewise, for each utterance with a falling contour, a matched utterance was digitally created with a rising contour. In this way, the same item could be heard with and without uptalk. Prolongation was treated as a between-item variable.

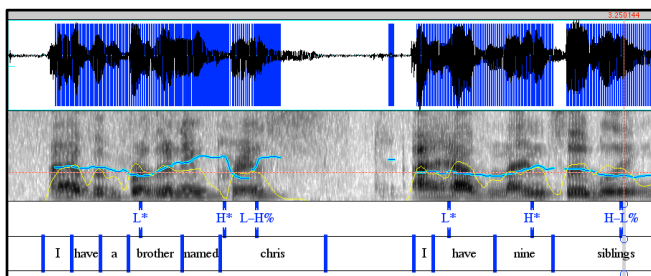


Figure 1: Example item.

**Design** Two lists were created so that each participant heard either the original or the manipulated version of an item. The lists were counterbalanced so that (1) original and manipulated versions were matched across lists, (2) half the items on each list were manipulated and half were not, and (3) rises and falls were equally likely to occur on a list.

**Procedure** Participants were told that the speakers they would listen to were recalling facts they had learned about celebrities, and that the speakers may have muddled their facts. Participants judged the likelihood that the speaker correctly reported the facts. The items were presented aurally with a screen displaying a 1 (not accurate) to 7 (accurate) Likert scale.

## Results

Here and in the subsequent two experiments, two-way repeated measures ANOVAs were conducted to test the effects of rising pitch and prolongations.

Items containing prolonged syllables at the end of the first utterance in the pair were rated as less accurate ( $M = 4.36$   $SE = 0.18$ ) than items with non-prolonged syllables ( $M = 4.94$   $SE = .21$ ;  $F(1, 19) = 14.92$ ,  $p < .001$ ). Items containing rising pitch at the end of the first utterance in the pair ( $M = 4.68$   $SE = 0.17$ ) were rated similarly to items with falling pitch ( $M = 4.62$ ,  $SE = 0.19$ ;  $F(1, 19) = .362$ ,  $p = .56$ ). There was no interaction,  $F(1, 19) = 1.90$ ,  $p = .31$ . See Table 1.

Table 1: Mean ratings for Experiment 1.

Condition	Means (SD)
Non-prolonged fall	4.96 (0.76)
Non-prolonged rise	4.91 (0.87)
Prolonged fall	4.28 (1.07)
Prolonged rise	4.44 (.87)

## Discussion

An off-line rating experiment testing listeners' assessments of speakers' knowledgeableness showed that prolongations decreased ratings of knowledgeableness. Surprisingly, rising pitch did not influence listeners' judgments as predicted by earlier studies.

One possible reason why these off-line ratings did not yield an effect of rising pitch is that listeners may have been unsure how to interpret the two hypothetical functions of rising pitch in this task. In Experiment 2, we test the predictions of the forward-looking and backward-looking hypotheses with an on-line task.

## Experiment 2

Listeners monitored for words that followed a sentence that ended with either a prolongation, uptalk, both a prolongation and uptalk, or neither a prolongation nor uptalk.

## Methods

**Participants** Twenty undergraduates participated in this experiment in exchange for course credit. There were 13 women and 7 men. All participants were native speakers of English.

**Materials** The same 48 stimuli were used as in Experiment 1. In addition, 48 filler stimuli were created. The filler stimuli were of two types. The first had the target word in the first utterance. The second did not contain the target word. The experiment contained half of each type of filler.

**Design** The design was similar to Experiment 1.

**Procedure** Each trial had the following structure. First, a 500 ms tone was heard indicating that the participants should focus their attention on the computer screen. The tone was followed by a 500 ms pause. A word appeared on the computer screen for 1000 ms, followed by a blank screen for 1000 ms, followed by the onset of the audio item. During the audio item, listeners pressed a button upon hearing the word presented visually previously on the screen. All target items and half of the fillers required a button press. The other half of the fillers required no button press. Targets were in variable positions in the second

utterances. Responses above 1500 ms and below 150 ms were excluded to eliminate false alarms and misses.

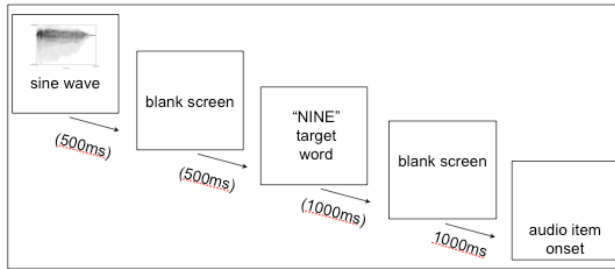


Figure 2: Trial presentation for Experiments 2 & 3.

## Results

Listeners monitored words faster after hearing prolonged syllables in the previous utterance ( $M = 550$   $SD = 93$ ) than non-prolonged syllables ( $M = 600$   $SD = 123$ ;  $F(1, 19) = 4.43, p < .05$ ). There was no effect of rising pitch ( $F(1, 19) = 0.7, p = .41$ ), but there was an interaction between prolongations and rising pitch ( $F(1, 19) = 5.31, p < .04$ ). Listeners monitored words fastest after prolonged rises ( $M = 525$   $SD = 87$ ) and slowest after non-prolonged rises ( $M = 653$   $SD = 158$ ). See Table 2.

Table 2: Mean reaction times for Experiment 2.

Condition	Means (SD)
Non-prolonged fall	546 ms (88 ms)
Non-prolonged rise	653 ms (158 ms)
Prolonged fall	575 ms (100 ms)
Prolonged rise	525 ms (87 ms)

## Discussion

Words were monitored for fastest when they followed sentences ending in prolonged rising pitch. Words were monitored for slowest when they followed sentences ending in non-prolonged rising pitch. These data show that prolongations can differentiate the two conflicting functions of rising pitch. Listeners gain a general processing advantage when comprehending material following prolongations. At the same time, prolongations interact with rising pitch in a way that makes prolonged uptalk more helpful for processing upcoming information and non-prolonged uptalk more harmful for processing upcoming information.

In a model with predetermined form-function relationships, these data would suggest that prolonged rising pitch equals forward-looking, and non-prolonged rising pitch equals backward-looking. But it may be the case that

listeners use beliefs about speakers' mental states in mapping intonational forms to different functions. In Experiment 3, we test how listener beliefs affect processing.

## Experiment 3

Listeners replicated Experiment 2 with new information about the speakers they would be hearing. About half were told that the speakers were non-experts who had effectively memorized the facts off cue cards and had trouble remembering them during the game. The other half were told that the speakers were knowledgeable experts on the facts they were saying.

## Methods

**Participants** Forty-two undergraduates participated in this experiment in exchange for course credit. There were 24 women and 18 men. All participants were native speakers of English.

**Materials** The same materials were used as in Experiment 2.

**Design** The design was similar to Experiments 1 and 2, except that Experiment 3 was effectively two separate experiments. In one, 22 participants were told that the speakers were non-experts in the fields they were talking about; that is, they knew little about pop culture, cinema, politics, and so on. In the other, 20 participants were told that the speakers were experts in the fields they were talking about. For example, they were led to believe that a person talking about politicians was a political science major and that a person talking about actors was a film major.

**Procedure** The procedure was the same as Experiment 2.

## Results

In the non-expert condition, there was no main effect of prolongations ( $F(1, 21) = 1.95, p = .17$ ) and no main effect for rising pitch ( $F(1, 21) = 0.44, p = .52$ ). There was, however, a significant interaction between prolongations and rising pitch ( $F(1, 21) = 9.85, p < .01$ ). These data are similar to those of Experiment 2, minus the effect of prolongations. See Table 3.

In the expert condition, there was also no main effect of prolongations ( $F(1, 19) = .01, p = .92$ ) and no main effect for rising pitch ( $F(1, 19) = 0.353, p = .56$ ). Unlike the non-expert condition, however, there was no interaction between prolongations and rising pitch ( $F(1, 19) = 1.87, p = .19$ ). See Table 4.

Table 3: Mean reaction times for Experiment 3 non-expert condition.

Condition	Means (SD)
Non-prolonged fall	521 ms (114 ms)
Non-prolonged rise	611 ms (138 ms)
Prolonged fall	570 ms (129 ms)
Prolonged rise	513 ms (115 ms)

Table 4: Mean reaction times for Experiment 3 expert condition.

Condition	Means (SD)
Non-prolonged fall	509 ms (113 ms)
Non-prolonged rise	550 ms (155 ms)
Prolonged fall	560 ms (187 ms)
Prolonged rise	506 ms (111 ms)

## Discussion

When listeners believed that speakers were non-experts, they interpreted rising pitch and prolongations similarly to when they were provided with no information about speakers (Experiment 2). However, when listeners thought that speakers were experts, their word monitoring was no longer affected by rising pitch or prolongations. This is supporting evidence that listeners establish relationships between linguistic form and function by first presupposing speakers' mental states.

### General Discussion

Temporal context, in this case prolongations, can affect off-line judgments of speaker knowledgeableness as well as how rising pitch influences on-line speech comprehension. At the same time, listener beliefs about speakers can affect whether prolongations and uptalk have any influence at all.

In Experiment 1, listeners considered utterances with prolongations as lacking in knowledgeableness. In contrast to expectation, uptalk had no effect on listener judgments. It may be that the off-line rating task was not sensitive enough to capture how listeners interpret rising pitch with the stimuli we used. In earlier studies, rising pitch indicated lower feeling of knowing with respect to the answers to questions. In our study, rising pitch was used with declarative statements. With declarative statements, there could be two interpretations for rising pitch, forward-looking and backward-looking.

In Experiment 2, prolongations gave listeners a processing advantage for monitoring words in the subsequent utterance. Although rising pitch alone had no effect, there was an interaction between rising pitch and prolongations. Prolonged rising pitch engendered the fastest reaction times for monitoring words in the subsequent utterance. Non-prolonged rising pitch engendered the slowest reaction times relative to all other conditions. This supports the hypothesis that prolonged rising pitch is forward-looking, and that non-prolonged rising pitch is backward-looking.

In Experiment 3, the interaction between rising pitch and prolongations was replicated for the group of listeners who believed that the speakers they were hearing did not really know what they were talking about. The interaction disappeared for another group of participants who believed that the speakers were experts on what they were talking about. Moreover, the data from Experiment 3 and from Arnold, Hudson Kam, & Tannenahus (2007) shows how inferences about speakers' mental states factors into the comprehension of Track 2 (Clark, 1996) or procedural (Wilson & Wharton, 2006) meaning. Where and how does this type of meaning arise? Our data would predict that this type of information is interpreted online during inferences about speakers' mental states.

While the lack of an effect of prolongations in Experiment 3 demonstrated that the findings in Experiment 2 were not driven by a confounding variable (for example, that prolonged items varied from non-prolonged items in a systematic way), we have no firm understanding of why the prolongation effect disappeared. Inspection of the means suggests that the Experiment 2 prolongation effect may have been driven by the much slower reaction times to the short rises. This suggests that listeners may have treated the doubt expressed in Experiment 2 by the short rises as more serious than the doubt expressed in Experiment 3 by the short rises. Evidence suggests that listeners assumed that speakers were not knowledgeable in Experiment 2 (as the results are similar to the non-expert condition in Experiment 3), but listeners were not explicitly told this. A short rise may have made listeners pay more attention to the utterance just produced in Experiment 2 because listeners were not primed to expect doubt (as they were in the non-expert condition of Experiment 3). Consequently, they may have taken the doubt more seriously and focused more on the utterance preceding the short rise.

The experiments presented here demonstrate that intonational events can be interpreted differently depending on temporal and situational context. Future research should examine how the functions of other intonational events might vary in different temporal and situational contexts.

### Acknowledgments

We would like to thank the following assiduous research assistants: Lisa Jenson, Sarah Jerz, Ben Radosevich, Shane Hihara, Jeannine Mesina, and Jack Emery.

## References

- Allen, K. (1984). The high rise terminal contour. *Australian Journal of Linguistics*, 4, 19-32.
- Arnold, J. E., Hudson Kam, C. L., & Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(5), 914-930.
- Beckman, M. E., & Ayers, G. M. (1994). Guidelines for ToBI Labeling.
- Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34(3), 383-398.
- Clark, H. H. (1996). *Using Language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H. (2002). Speaking in Time. *Speech Communication*, 36, 5-13.
- Clark, H. H., & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speech. *Cognition*, 84, 73-111.
- Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.
- Fletcher, J., Stirling, L., Mushin, I., & Wales, R. (2002). Intonational rises and dialog acts in the Australian English Map Task. *Language and Speech*, 45(3), 229-252.
- Fox Tree, J. E. (2001). Listeners' uses of *um* and *uh* in speech comprehension. *Memory and Cognition*, 29(2), 320-326.
- Gunlogson, C. A. (2001). *True to Form: Rising and Falling Declaratives as Questions in English*. Unpublished Doctoral Dissertation, University of California Santa Cruz.
- House, J. (2007). The role of prosody in constraining context selection: A procedural approach. *Nouveaux cahiers de linguistique française*, 28, 369-383.
- Kohler, K. J. (2004). *Prosody revisited: Function, time, and the listener in intonational phonology*. Paper presented at the Proc. International Conference: Speech Prosody 2004, Nara, Japan.
- McLemore, C. A. (1991). *The Pragmatic Interpretation of English Intonation: Sonority Speech*. Unpublished Doctoral Dissertation, University of Texas at Austin.
- Murray, I. R. & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, 93(2), 1097-1108.
- Pierrehumbert, J., & Hirschberg, J. (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. In P. R. Cohen, J. Morgen, & M. E. Pollack (Eds.), *Intentions in Communications* (pp. 271-311). Cambridge, MA: The MIT Press.
- Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech synthesis. *Journal of the Acoustic Society of America*, 105(1), 512-521.
- Shokier, V. (2008). Evidence for the stable use of uptalk in South Ontraio English. *Working Papers in UPenn Linguistics*, 14(2), pgs. 16-21.
- Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, 38(10), 1559-1579.