# The form and use of uptalk in Southern Californian English

*Amanda Ritchart[1], Amalia Arvaniti[2]*

[1] University of California, San Diego
[2] University of Kent
aritchart@ucsd.edu, a.arvaniti@kent.ac.uk

## Abstract

This study examines the phonetics, phonology and pragmatic function of uptalk, utterance-final rising pitch movements, as used in Southern Californian English. Twelve female and eleven male speakers were recorded in a variety of tasks. Instances of uptalk were coded for discourse function (statement, question, confirmation request, floor holding) based on context. The excursion of the pitch rise and the distance of the rise start from the onset of the utterance's last stressed vowel were also measured. Confirmation requests and floor holding showed variable realization. Questions, on the other hand, showed a rise that typically started within the stressed vowel and had a large pitch excursion, while uptalk "proper", i.e. uptalk used with statements, exhibited both a smaller pitch excursion and a later rise that often started after vowel offset. This pattern suggests that statements have a L* L-H% melody while questions have L* H-H%. Gender differences were also found: female speakers used uptalk more often than males, and showed greater pitch excursion and later alignment, all else being equal. Other social parameters, however, such as social class and linguistic background, did not affect the use of uptalk.

**Index Terms**: intonation, HRT, uptalk, English, sociolinguistics, gender

## 1. Introduction

Rising melodies used with statements, commonly referred to as *uptalk* or *high rise terminals* are common in many varieties of English. Here we use the term *uptalk* which better reflects the Southern Californian patterns that are the focus of our investigation. Research on uptalk in some varieties is quite extensive, but has often been impressionistic [1]. The varieties that have been most investigated include those spoken in Australia and New Zealand as well as UK varieties from Glasgow and Belfast [1]–[5] (and references therein).

These studies document that different tunes are used for uptalk across varieties. Thus, [1] report that Australian uptalk is realized as either L* H-H% or H* H-H%. For Glasgow, L*H H-L% is proposed for the "rise-plateau-slump" type of uptalk, with suspension of the rule that in other English varieties upsteps a L% after a H- phrase accent [2]. In [1], New Zealand uptalk is analyzed as reflecting two main patterns, LH* H-H% and L* H-H% (based on [3]), but a newer study suggests that New Zealand English may exhibit change in progress with respect to uptalk [4].

In addition to differences in form, uptalk across varieties of English is used for different purposes. Thus, [1] report that in Australian English upstep is used both with questions and declarative statements; upstepped statements are particularly frequent when the speaker wishes to hold the floor. This leads [1] to suggest that the intonational difference between statements and questions and that between statements and continuation is neutralized in Australian English. New Zealand

English also uses uptalk for both statements and questions but the tunes used for each function are becoming increasingly distinct [4]. Research on Glasgow and Belfast English, e.g., [2] and [5], focused on form rather than function, but recent research suggests that uptalk, in Belfast at least, may have its origins in list intonation [6].

One of the varieties that is stereotypically known as exhibiting use of uptalk is Californian English, particularly the varieties spoken in the south (henceforth *SoCal*). The use of uptalk in SoCal is often referred to as "valley girl speak" and is often assumed to be a feature of younger females only, though no studies exist, to our knowledge, confirming or refuting this general lay perception.

Here we present data from SoCal English which show that the use of uptalk is widespread in this variety and exhibits gender-related variation. We further show that SoCal uptalk tunes are different from those reported for other varieties of English, and that speakers retain systematic differences between uptalk used in statements and other types of uptalk, such as pitch rises used with questions. Differences apply both to the tunes employed and to the scaling of the rise.

## 2. Methods

### 2.1. Speakers

Twenty-three speakers were recorded for the study, eleven male and twelve female. They were all native speakers of SoCal English, from San Diego (N = 7), Orange (N = 6), Los Angeles (N = 8), and Riverside (N = 2) counties. Fifteen were monolingual, while the other eight reported being bilingual in English and one of the following languages: Vietnamese (N = 3), Japanese (N = 1), Armenian (N = 1), Assyrian (N = 1), Spanish (N = 1), and Cantonese (N = 1). The speakers' ethnic backgrounds varied: twelve self-identified as Asian, six as Hispanic and five as White.

The MacArthur Scale of Subjective Social Status (cf. [7], [8]) was used to determine the speakers' socioeconomic status or *SEC*, a rather fluid concept in California. Participants found the use of the scale easy and intuitive. They were classed into three groups based on their responses: lower (rungs 1-4, N = 4), middle (rungs 5-7, N = 13) and upper (rungs 7-10, N = 6).

### 2.2. Materials, Tasks and Procedures

Recordings took place in the recording studio of the UCSD Phonetics Lab, using an AD converter at 48 KHz and 16-bit quantization. Four types of data were collected from each speaker: (a) map task; (b1) reading of the transcript of a popular sitcom scene; (b2) retelling of the sitcom scene; (c) controlled materials consisting of isolated questions and statements. For the first 17 participants, tasks were presented in the following order: (c), (b), (a). (For task (b), the retelling of the clip always followed the reading of the transcript.) To control for possible order effects, the order of tasks was

counterbalanced for the other recordings and each participant was randomly assigned to one of three possible orders (Latin square design): abc, bca, or cab. However, given that the data consist largely of spontaneous speech it is unlikely that order could have severely biased speaker productions with respect to uptalk.

For the map task, maps with local (or local sounding) landmarks were designed as illustrated in Figure 1. In the map task, the participants acted as leaders with the follower being either the first author or an undergraduate research assistant (both females and native SoCal speakers). For task (b1), a scene from either *Scrubs* or *How I met your mother* was used; the show chosen was the one the participant was less familiar with. Lack of familiarity was sought so that speakers would not imitate the actors' accents. The scene was muted and participants were given a transcript of the dialogue while they watched the clip. When they were ready, they chose which character they were most comfortable reading from, and participated in reading aloud the transcript in a dialogue with the experimenter. In task (b2), participants had to retell the same scene in their own words. For task (c), the participants read aloud a list of 49 sentences. These were statements and questions constructed for the study. In these sentences, the number of syllables and position of stress was controlled in order to examine the realization of specific tonal events; see (1) for an illustration. Here we report on the results from tasks (a) and (b2).

(1)   a. *Did Anne and Mel eat the lime?*
        b. *Did Annabelle and Melinda eat the lime?*



Figure 1: *The Instruction Giver's map used in the map task.*

## 2.3.  Analysis and Measurements

The analysis involved both a categorization of each instance of uptalk in terms of its discourse function and acoustic measurements with respect to the alignment and scaling of the pitch rise associated with each uptalk token.

Specifically, instances of uptalk were classed in one of four discourse functions: question, statement, holding the floor, and confirmation request. A token was considered to be a question when it was syntactically marked as such, for example by showing inversion. Confirmation requests were indirect questions: they were not syntactically questions, but the context and interlocutor response indicated that the speaker was indirectly asking if their interlocutor was paying attention, agreed or understood. Holding the floor was defined as an utterance indicating that the speaker did not intend to cede the

floor, in that s/he continued talking with either a minimal or no pause and was not interrupted by their interlocutor. All other instances of uptalk were identified as statements. These were regular declaratives for which no other discourse function was apparent from context; e.g., such utterances did not elicit information from the interlocutor. For cases in which the discourse context was ambiguous, a forced choice was made by the first author who is a native speaker of the dialect.

In addition, the scaling and alignment of the rise was annotated using the facilities of Praat [9]. The beginning of the rise was manually located as the point at which an upward trend was apparent in which successive F0 values differed by more than 5 Hz (this was done to exclude microprosodic variation); see Figure 2 for an illustration. Scaling was measured in Hz and defined as the difference between the F0 at the beginning of the rise and the highest F0 point at its end. After the F0 information was extracted, values in Hz were converted to ERB in order to better compare male and female voices. The alignment of the F0 rise was defined as the distance of the point annotated as the start of the rise from the onset of the last stressed vowel in the utterance. This measurement was based on the assumption (supported by the data) that the last content word is typically the one carrying the nuclear pitch accent.

## 3.   Results

Results presented here are related to the function, scaling and alignment of rises and to differences in gender. We note that ethnicity, SEC status and bilingualism did not affect the use of uptalk; thus they will not be discussed further. All significance testing was determined using linear mixed-effects models with Speaker as a random intercept. P-values are given with respect to model comparisons and are reported with the $\chi^2$ statistic, which compares the (reduced) model without the fixed effect and the (full) model with the fixed effect.

### 3.1.  Discourse Functions and Distribution of Uptalk

The best-fit model for comparing uptalk against other utterances in the corpus included discourse function, task type, gender and an interaction between gender and discourse function as fixed effects. Uptalk was more frequent in the map task than in clip retell: 34% of the utterances in the map task ended in uptalk as opposed to only 20% of utterances in clip retell [$\chi^2(1) = 37.4$, $p < 0.001$]. Uptalk was also used more frequently, approximately twice as often, by female than male speakers: uptalk comprised 42% of the female speakers' utterances vs. 20% of the male speakers' utterances [$\chi^2(1) = 14.1$, $p < 0.001$].

Gender also interacted with discourse function [$\chi^2(3) = 16.9$, $p < 0.001$]. First, no gender or discourse differences were found for uptalk in questions and confirmation requests in the corpus: uptalk was used for both types of utterances in 100% of the tokens independently of speaker gender. Statements and floor holding, on the other hand, showed different frequencies for uptalk, with floor holding being signaled by uptalk significantly more frequently than statements: 45% of floor holding ended in uptalk vs. 16% of statements [$\chi^2(3) = 244.7$, $p < 0.001$]. However, while females and males used uptalk with statements equally frequently, females used uptalk to hold the floor significantly more frequently than males; indeed females used uptalk more than twice as much as males for floor holding. This is illustrated in Figure 3.
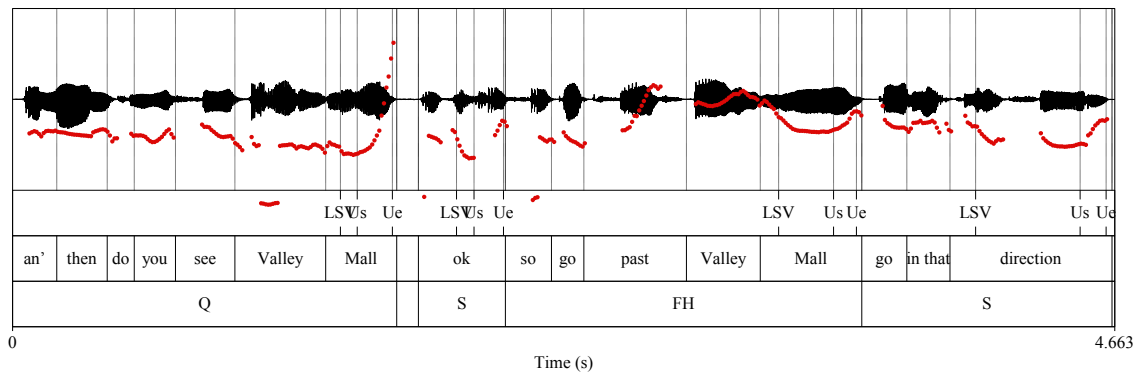
Figure 2: *Example of data annotation from the map task. LSV = last stressed vowel; Us = start of uptalk rise; Ue = end of uptalk rise; Q = question; S = statement; FH = floor holding. The follower's response ("yes, I do") which followed the question in this example has been removed for clarity.*
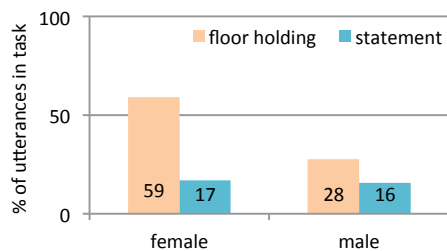


Figure 3: *Proportion of uptalk used by discourse function and gender.*

### 3.2. Alignment of Uptalk Rise

The best-fit model for the alignment of the uptalk rise included discourse function and gender as fixed effects. In this model, only two levels of discourse function were included, statement and question. Floor holding and confirmation requests were omitted from the model as their alignment was too variable.

The results from statements and questions showed a consistent difference between the onset of the rise in statements vs. questions, with the former having significantly later alignment than the latter [$\chi^2(1) = 19.3$, $p < 0.001$]. Specifically, the rise in the questions included the last stressed vowel (which is presumed to carry the nuclear pitch accent) while in statements the rise started after this vowel. The difference in the alignment of the rise in statements and questions is illustrated in Figure 4, which also shows the effect of gender on alignment. Specifically, uptalk produced by female speakers showed later alignment than uptalk produced by male speakers both for statements and questions [$\chi^2(1) = 5.6$, $p = 0.02$]. The differences were quite substantial, particularly for the questions: male speakers started the rise just before the last stressed vowel on average, while the rise for the female speakers started within this vowel.

### 3.3. Scaling of the Uptalk Rise

The best-fit model for the scaling of the rise included discourse function [$\chi^2(3) = 19.4$, $p < 0.001$], gender [$\chi^2(1) = 27.01$, $p < 0.001$] and task type [$\chi^2(1) = 20.03$, $p < 0.001$] as fixed effects. The major difference in pitch excursion with

respect to discourse function was between statements and the other functions, with statements showing approximately half the pitch rise than questions, confirmation requests and floor holding (see Figure 5). Differences between these last three discourse functions were also statistically significant but minimal in actual terms [questions, confirmation requests > floor holding].
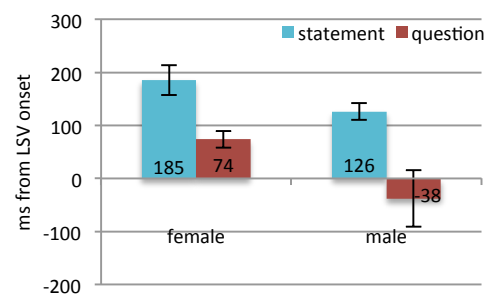


Figure 4: *Mean rise alignment (with standard error bars) per type of discourse function and gender. Negative values represent a rise beginning before the onset of the last stressed vowel (LSV).*
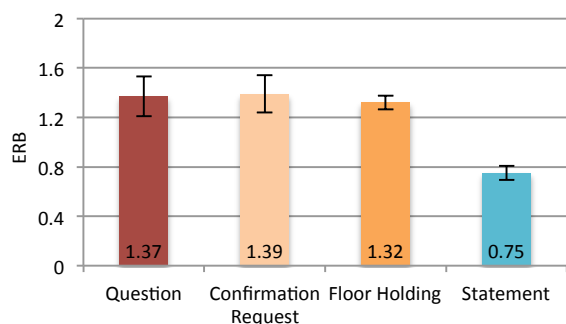


Figure 5: *Mean scaling of rises (with standard error bars) per discourse function.*

In addition, the data showed that female speakers had generally greater pitch excursions than males (see Figure 6a), presumably a reflection of gender differences in the use of the

frequency and effort codes [10]. Further, pitch excursions associated with uptalk were significantly larger in the map task than in clip retell (see Figure 6b). Neither result interacted with discourse function, however, suggesting these are independent effects and not the result of, e.g., female talkers asking more questions, or speakers in general making more confirmation requests in the map task than in clip retell.
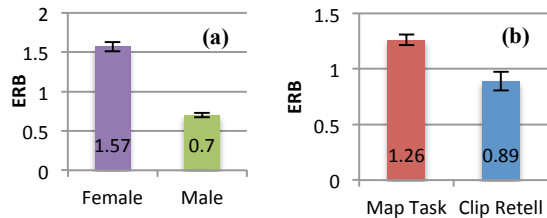


Figure 6: *On the left, mean scaling of uptalk (with standard error bars) by gender; on the right, mean scaling of uptalk (with standard error bars) by type of task.*

## 4. Discussion and Conclusions

Given the above results, we propose that the melody *typically* used with questions in SoCal English is L* H-H% and that used with statements is L* L-H%. The difference in phonological composition accounts both for the difference in alignment reported above but also for the difference in the scaling of the pitch rise: L-H% results in a lower rise than H-H%. Questions, as noted, can show a rise on the stressed syllable, a contour that could be interpreted as the reflex of a bitonal LH accent. However, the auditory impression is that of a low pitch accent, while the use of either L*H or LH* in questions is pragmatically doubtful (cf. [11] on the pragmatics of L*H when followed by a rise). Independently of the representation adopted for the question tune, the fact remains that questions and statements are not relying on the same melody as is often assumed; to put it differently, SoCal statements with uptalk do not sound like questions.

Our results further show that SoCal English makes a distinction between uptalked statements and questions even when the same melody is used (as happens occasionally). In particular, although the distinction is typically realized as a choice of tune, as noted above, it can also be signaled by just differences in the pitch scaling of the final rise (cf. the question and statements in Figure 2). The difference in pitch rise scaling is particularly evident when the tune used is H* H-H%, a variant that was attested but was not as frequent in our data as the L* accent variants. If such differences in the scaling of the rise turn out to be used by listeners to interpret the pragmatic intent of an utterance, this would suggest the need to incorporate scaling contrasts beyond H vs. L in phonological representations of intonation.

The two main melodies L* H-H% and L* L-H% are also used for floor holding and confirmation requests except that these two functions do not have as consistent a connection with a specific melody. In the case of confirmation requests this could be due to their dual role as questions and statements: speakers are making a statement but simultaneously requesting that their interlocutor confirm that what is said is understood or accepted. Thus, speakers use L* H-H%, L* L-H% or H* H-H% in these instances. Regarding floor holding, one of the most noticeable features was the use of high plateaux, rather than rises per se. Plateaux are particularly prevalent when speakers are listing items or instructions in the map task (cf.

[12] on the intonation of lists). Plateaux are possible realizations of high tones [13] and thus they can perhaps create the impression of a rise; however, in our data they were clearly different from uptalk "proper" both acoustically and impressionistically and thus best represented phonologically as L* H-L% where the L% is upstepped.

The patterns described above document the use of tunes that are different from those described for other varieties of English that use uptalk. In particular, the prevalence of L* is not reported for other varieties of English (but see [4] on New Zealand English). As noted, for example, Australian English uses mostly a H* accent and it is precisely this use that has given rise to the term *High Rising Terminal*. Thus, the present study underlies the importance of including dialectal variation in the investigation of intonation and gives support to the claim that such variation exists even within dialectal areas often described as uniform, like the USA West [14].

Regarding the demographic factors in our study, we note that there are consistent differences between genders, with females using uptalk twice as often as males. This difference is presumably what has given rise to the stereotype that uptalk is used by females only; among women uptalk is sufficiently frequent to be identified as a distinctive characteristic of their way of speaking. Contrary to the popular stereotype, no gender differences in the use of uptalk were observed for statements: approximately 16% of statements ended in uptalk in the speech of both men and women in our sample. However, differences are evident in the use of uptalk for floor holding: in this use, uptalk is twice as frequent in the data from female speakers (a result that in itself suggests that the similar frequency of uptalk with statements cannot be attributed to the fact that our male speakers interacted with female researchers). At present we do not have a good explanation for this but offer some suggestions. One possibility is that women wish to hold the floor longer and use uptalk as a device to indicate this intent. This explanation however does not quite tally with existing research suggesting that women do not take longer turns than men ([15] and references therein). Another possibility is that women wish to indicate their intent to hold the floor because they are generally interrupted more often than men [15]. Again, this is not entirely satisfactory as our data were based on monologues (clip retell) and a cooperative task in which the interlocutor was always female. Thus, this aspect of the data clearly requires further investigation. At the same time, we do find that gendered use of uptalk did not interact with task. From this we can infer that the gender effect is not due, e.g., to women asking more questions, but rather to their general preference for certain uses of uptalk.

Unlike gender, which was a clear determiner of the frequency, function and form of uptalk, we did not find differences relating to ethnicity, SEC status or the language background of our speakers. Although it is possible that such differences could emerge with a larger sample, the ubiquitous use of uptalk in our corpus rather suggests that uptalk is sufficiently widespread in SoCal to transcend social barriers. In turn this tallies with the speakers' attitude to uptalk: for SoCal speakers it is not a feature that attracts attention.

## 5. Acknowledgements

# 6.  References

[1]   J. Fletcher, E. Grabe, and P. Warren. "Intonational variation in four dialects of English: the high rising tune," in *Prosodic Typology: The Phonology of Intonation and Phrasing*, S-A. Jun, Ed. Oxford: Oxford University Press, 2005, pp. 390-409.

[2]   C. Mayo, M. Aylett, and D. R. Ladd. "GlaToBI prosodic transcription of Glasgow English: An evaluation study of GlaToBI," in *Proceedings of the ESCA Workshop on Intonation: Theory, Models and Applications*, 1997, pp. 231-234.

[3]   N. Daly and P. Warren, "Pitching it differently in New Zealand English: Speaker sex and intonation patterns," *Journal of Sociolinguistics*, vol. 5, no. 1, pp. 85-96, 2001.

[4]   P. Warren, "Patterns of late rising in New Zealand: Intonational variation or intonational change?" *Language Variation and Change*, vol. 17, no. 2, pp. 209-230, 2005.

[5]   E. Jarman and A. Cruttenden, "Belfast intonation and the myth of the fall," *Journal of the International Phonetic Association*, vol. 6, no. 1, pp. 4-12, 1976.

[6]   J. Sullivan. "The why of Belfast rises," in *New Perspectives on Irish English,* B. Migge and M. Ní Chiosáin, Eds. John Benjamins Publishing Company, 2012, pp. 67–84.

[7]   N. E. Adler, E. S. Epel, G. Castellazzo, and J. R. Ickovics, "Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women," *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, vol. 19, no. 6, pp. 586-592, 2000.

[8]   A. Singh-Manoux, M. G. Marmot, and N. E. Adler, "Does subjective social status predict health and change in health status better than objective status?" *Psychosomatic Medicine*, vol. 67, no. 6, pp. 855-861, 2001.

[9]   P. Boersma and D. Weenik. (2013). *Praat: doing phonetics by computer* [Computer program], Version 5.3.59. Available: http://www.praat.org

[10]  C. Gussenhoven, *The Phonology of Tone and Intonation*. Cambridge University Press, 2004.

[11]  J. Hirschberg and G. Ward, "The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English," *Journal of Phonetics*, vol. 20, pp. 241-251, 1992.

[12]  M. Liberman and J. Pierrehumbert. "Intonational invariance under changes in pitch range and length," in *Language Sound Structure*, M. Aronoff and R. Oehrle, Eds.  Cambridge, MA: MIT Press, 1984, pp. 157-233.

[13]  R-A. Knight and F. Nolan, "The effect of pitch span on intonational plateaux," *Journal of the International Phonetic Association,* vol. 36, no. 1, pp. 21-38, 2006.

[14]  W. Labov. "The three dialects of English," in *Handbook of Dialects and Language Variation*, M. D. Linn, Ed. San Diego: Academic Press, 1998, pp. 39-81.

[15]  J. Coates, *Men, women and language*. Pearson Education, 2004.