

Sociophonetic and prosodic influences on judgements of sentence type

Paul Warren

Victoria University of Wellington

paul.warren@vuw.ac.nz

Abstract

A mouse-tracking experiment is presented in which utterances containing different realisations of the SQUARE vowel and early or late intonational rises are used in a forced-choice decision task. The selection of Question or Statement responses is affected by the detail of the rising intonation, as is the speed of the response. Track data from mouse movements provides further qualitative indication that response confidence depends not just on rise alignment but also on the nature of the vowel, reflecting the merger-in-progress of NEAR and SQUARE in New Zealand English.

Index Terms: uptalk, vowel merger, sentence type, New Zealand English, mouse-tracking

1. Introduction

The distinction between falling and rising intonation is typically but erroneously linked to the distinction between statements and questions. While it is possibly true that the only intonational universal is that high pitch indicates ‘open’ and low pitch indicates ‘closed’ [1], it is also true that many questions have falling intonation. There is a negative correlation between the amount of morphosyntactic marking of questionhood in English and the likelihood of rising intonation [2, 3]. For instance, *wh*-questions have both the *wh*-word and subject-auxiliary inversion, and are more likely to have falling than rising intonation.

In addition, of course, there are statements that have rising intonation. As well as list intonation and continuation rises, New Zealand English (NZE), like many other varieties, is also noted for the presence of uptalk rises, often referred to as high rising terminals. These are rising intonation patterns on utterances that are intended as declarative statements. As is often found for speech phenomena that are undergoing change, and which are therefore more typical of younger speakers, uptalk has been subjected to some bad publicity, with labels such as ‘moronic interrogative’ or ‘idiotic-sounding antipodean rise’. Because of its link with questioning, the rise is often interpreted in media reports and in letters-to-the-editor as indicating that the speaker is uncertain of what they are saying, or that they are at the very least deferring to their interlocutor. The more positive interpretation found in the linguistic literature is that uptalk signals inclusion of the interlocutor (another trait of the ‘openness’ of high pitch) and a more collaborative speaking style. (For a summary of the literature on uptalk in English and other languages, see [4].)

Research on uptalk in NZE has focused on two areas. One is the types of discourse structures that it occurs in – these are typically narratives and descriptive texts, rather than opinion texts, a finding that is seen to undermine the claim that uptalk is an indication of uncertainty or deference. If it conveys uncertainty then we would not expect it on narratives and

descriptions, and if it shows deference then we would expect to find it more in opinion texts [5]. The other focus has been on the demographic characteristics of uptalkers. These are typically younger females, with a higher incidence also amongst Māori than Pākehā.

NZE speakers not only use rises on many of their statements, but of course they also continue to use them, as do other speakers, on questions. This raises the issue of whether there are any differences between the types of rises in the statements and questions from these speakers. Research has suggested that there are, and that there may be a change in progress – the starting point for the rise in questions is becoming earlier, relative to the accented syllable, for younger NZE speakers. This distinguishes question rises from uptalk rises, which are later and more dramatic (see Figure 1) [6]. The ability of listeners to utilize this distinction has been tested in an off-line forced choice (‘question’ or ‘statement’) task using an utterance (*He’ll be at the basketball stadium*) with manipulated pitch contours that had rises starting in five different positions from the accented syllable *ba-* through to the final syllable *-um* [6, 7]. This study found that earlier rises were more likely to result in ‘question’ choices. The study used two young speakers, a male and a female, and reported no significant difference between the two.

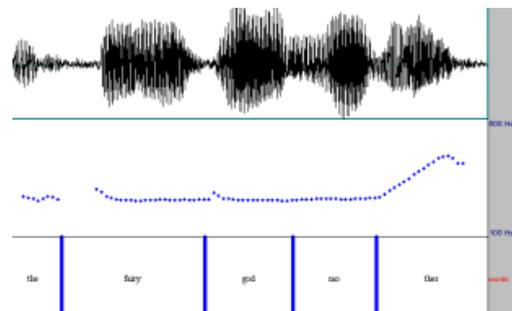


Figure 1 – Late uptalk rise on the phrase ‘fairy godmother’ from a NZE speaker

2. Sociophonetic cueing

Given the stronger tendency for uptalk from younger speakers, it is relevant to conjecture whether other cues to the speaker’s age might have an impact on the interpretation of rising intonation patterns, in particular on the distinction between early rises for questions and late rises for statements. Previous perceptual research has manipulated cues such as the age of a person – purportedly the speaker – in a photograph presented before the speech stimulus. After a photograph of an older person, participants were more accurate in reporting the identity of a word stimulus that contained one of two vowels that are still distinct in older and more conservative speakers but are becoming merged in the speech of younger New

Zealanders [8]. This vowel pair is NEAR and SQUARE, as in *beer* and *bear*, *cheer* and *chair*, etc.

The current study exploited this merger of NEAR and SQUARE to provide a sociophonetic cue that might influence listeners' perceptions of the likely age of the speaker, and thus influence their interpretation of rising intonation patterns on the utterance as indicating either a question or a statement. This allows us to measure the impact of speaker characteristics on the interpretation of rises without directly drawing attention to the age of the speaker. Sociophonetic research indicates that the NEAR-SQUARE merger is a good candidate for this investigation, since a merger in the direction of the NEAR vowel is well established for younger speakers but not for older speakers [9]. The age stratification of the merger parallels that of the incidence of uptalk. Both are more frequent for younger speakers. So too is the distinction between early and late rise onsets for questions and statements.

3. Experiment

The experiment reported here investigated two hypotheses. Firstly, that early and late rises on utterances with declarative word order will correspond to interpretations of the utterances as questions vs. statements respectively. Secondly, that the presence of a closer, NEAR-LIKE realisation of a SQUARE vowel in an utterance will indicate a younger speaker, resulting in strengthening of the first hypothesis for such an utterance.

3.1. Materials

To test these questions, utterances were devised that contained a word that contained a SQUARE vowel (such as *cared*) which was manipulated so that it had either the more open starting point appropriate for SQUARE or the closer starting point of NEAR. These words appeared in the first half of the utterance. The end of the utterance had a rising pitch contour that was manipulated so that the rise had either an early starting point (at the beginning of the accented syllable, e.g. *animals*) or a later starting point (on the final syllable, e.g. *animals*). An example sentence is *John's mother cared for stray animals*.

Twenty such utterances were used, each of which had four manipulations, crossing NEAR or SQUARE against early or late rises. A young female speaker of NZE recorded two versions of each utterance, intending them – as instructed – either as questions or as uptalk statements. For 10 of the test utterances the question version was selected, and for the other ten the statement version. The manipulation of the vowel in the square word (e.g. *cared*) was manipulated using a Praat script written by the author for manipulation of formant frequencies. The final rise was also manipulated in Praat. Rises climbed from around 200Hz to around 400Hz, either from the onset of the voiced portion of the accented syllable or from the onset of the voiced portion of the final syllable. The rise was of the same magnitude for the early and late rise versions of each utterance.

In addition, the same speaker recorded 20 statements with falling intonation patterns, 10 questions with inversion and rising intonation and 10 wh-questions with falling intonation. None of these filler sentences contained any words with either NEAR or SQUARE. Two test lists were made, one containing the SQUARE vowel versions of SQUARE words and the other containing the NEAR vowel versions. Each list contained both early and late rise versions of each utterance, in two blocks,

with equal numbers of each rise type in each block, and with no item repeated within a block. Each test list contained all 40 filler items, half of which were repeated during the test, so that participants experienced some clearly repeated items as well as the potentially repeated test items (the same utterances with early and late rises). Thirty of the resulting 60 fillers were in each block. Each test list was therefore made up of two blocks of 50 items each.

3.2. Methodology

The method used for this experiment was a forced choice task, where the choices were 'statement' or 'question'. In order to obtain information about the confidence of participants' choices, the mouse-tracking procedure was chosen. Participants sit in front of a computer screen and click on a box at center bottom of the screen marked 'start'. They then hear a stimulus, and have to indicate their choice by clicking on one of two boxes, marked 'question' or 'statement' in the top corners of the screen (see Figure 2). The track taken by the mouse in getting to those responses is recorded. Responses were timed out 5 seconds from the beginning of the utterance. Half of the participants assigned to each test list saw 'question' in the top left box and 'statement' in the top right box. The other half had these positions reversed.

In connection with mouse-tracking it has been claimed that 'hand in motion reveals mind in motion' [10], i.e. that the track followed in getting from the 'start' button to one of the response buttons gives an indication of the manner in which the participant reaches that decision. A direct movement from 'start' to the selected response indicates a more confident response than one that moves first towards one response button before then selecting the other.

A number of measures are available from this task. First, there are the standard measures found in much behavioural research of the choice made and the time taken to make that choice. Then there are the specific measures obtained by tracking the mouse positions. Two of these are measures of spatial attraction, as illustrated in Figure 2 – Area Under the Curve (AUC) and Maximum Deviation (MD). AUC is an index of overall attraction towards the unselected alternative (incorporating all time steps recorded in the trial). MD indicates maximum attraction to the alternative.

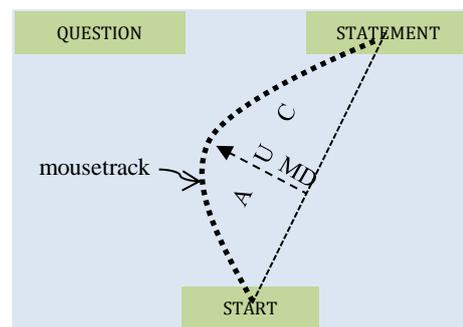


Figure 2: Example of mouse-tracking measures

A mouse-tracking procedure was scripted by the author in E-Prime [11]. Post-processing of the track data was carried out in the R language [12], using scripts written for this purpose by the author.

3.3. Participants

Participants were 20 native speakers of NZE aged between 18 and 38; 13 were females. All were right-handed mouse users. Participants were allocated randomly and evenly to the two test lists described above.

4. Results

The error rate for filler statements (responding that a statement was a question) was very low at 0.26%. For filler questions it was higher at 3.84%. Fillers were not further analysed.

The data from the test utterances were analysed in mixed effects models. Fixed effects included the participant variables Sex and Age, and the experimental variables Vowel (NEAR or SQUARE), Rise (early or late) and the Sequential position of the stimulus in the experiment. By-participant random slopes of sequential position were included where model comparison showed their inclusion to improve the model fit.

4.1. Response choice

Of 800 possible responses to test items, 43 (5.4%) were missing either for technical reasons or because the participant did not respond within the 5s timeout period. Overall there were more Question than Statement responses (74.5% Question responses). Participants' response choice was coded as 0 for Question and 1 as Statement, and entered into a logistic regression model. Response choice was affected neither by participants' Age nor by their Sex. Including Sequential position as a random slope improved the model fit, but Sequential position itself was not a significant predictor of response choice, nor did it interact with Rise (indicating that there was no shift in response patterns to the different Rise shapes with increased exposure during the experiment).

Turning to the experimental factors, there was no significant effect of Vowel, but there was for Rise, with more Question responses after early rises (80.6%) than after late rises (68.4%; $\beta=1.29$, $SE=0.25$, $p<0.001$). This result matches findings in earlier production and perception studies [6] and supports the first hypothesis set out above. However, the absence of a significant interaction between Vowel and Rise means that there is no evidence that presence of a NEAR realisation of a SQUARE vowel increases the likelihood that an early rise will be interpreted as a question and a late rise as a statement, and so the second hypothesis is not supported.

For the remaining analyses, only Question responses are included, representing the majority response choice. There were too few 'Statement' responses across the conditions of the experiment for reliable statistical analysis. An additional factor included in these analyses is Box, a variable indicating whether the Question response was made in the response box on the left or on the right.

4.2. Response times

Because of a skewed distribution typically found with response times (RTs), these were log-transformed before being entered as the dependent variable in linear regression modelling. In the final best-fit model, Sex, Age and Box were not significant. There was a marginal effect of Sequence, indicating an unsurprising trend for responses to get faster as the experiment progressed ($t=-1.91$, $p<0.06$). Vowel was significant neither as a simple effect nor in interaction with Rise. There was a significant simple effect of Rise: responses after early rises (2435msec) were faster than those after late

rises (2584msec; $t=3.22$, $p<0.02$). This pattern of results strongly supports the hypothesis that early rises are more strongly associated than late rises with question utterances.

4.3. Area Under the Curve

Mixed effects regression modelling returned significant simple effects for Sequence ($t=-4.70$, $p<0.001$), with responses tracking more directly to the response box as the experiment progressed, and for Box ($t=3.05$, $p<0.02$), reflecting the fact that responses were more direct when the Question option was on the left. This probably reflects physiological differences involved in moving the mouse to the two response boxes. There was an interaction effect of Rise with Vowel ($t=2.18$, $p<0.05$): after a NEAR vowel, early rises resulted in more direct Question responses than late rises; after a SQUARE vowel there was no significant difference in responses for early and late rise stimuli (see Figure 3).

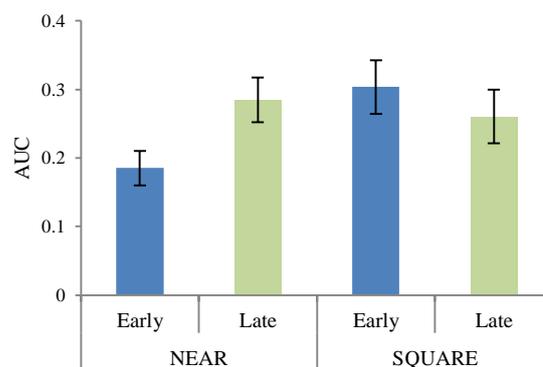


Figure 3: Area Under the Curve by Vowel and Rise, for Question responses

4.4. Maximal Distance

In the mixed effects analysis of Maximal Distance, we again find a significant effect for Sequence ($t=-3.46$, $p<0.001$), and Box ($t=2.64$, $p<0.05$). The patterns for these factors are the same as for AUC. Again, there was a significant interaction of Rise with Vowel ($t=2.44$, $p<0.02$; see Figure 4), with a clear effect of Rise after NEAR but not after SQUARE.

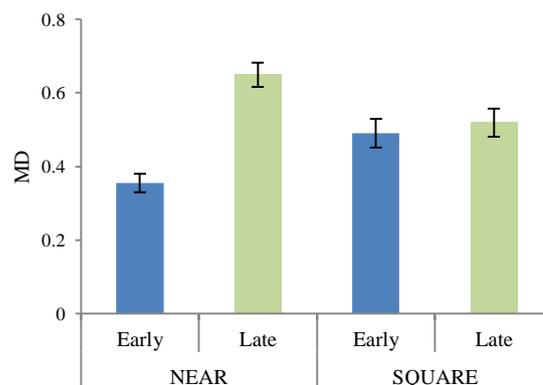


Figure 4: Maximal Distance by Vowel and Rise, for Question responses

5. Discussion

The results presented above provide clear support for the first of our two hypothesis, and indirect support for the second. The first hypothesis was that on utterances with declarative word order, such as *John's mother cared for stray animals*, an early intonation rise would indicate a question and a late rise a statement. Recall that early and late are relative terms here – both rises started on the final word of the utterance, either on the accented syllable (the first syllable in the example) or on the final syllable. The response required in the mouse-tracking task was a binary forced-choice decision between Question and Statement. The analysis of response choices showed a clear shift of the pattern of responses, with an increase in Question responses when the rise started early, compared with when it started late. The hypothesis is also supported by the more detailed analysis of response times in the majority Question responses, which showed that these were faster after early rises.

The second hypothesis was that the realisation of a particular vowel in the lead-up to the rise would have an influence on the interpretation of that rise. The vowel in question was a SQUARE vowel (in the word *cared* in the example above), which was manipulated in different versions of the test utterances to be more SQUARE-like or more NEAR-like. Since NEAR realisations of this vowel are more typical of younger speakers of NZE, it was predicted that the presence of the NEAR-like vowel would make it more likely that the early rise would be associated with a question and the late rise with a statement, since this pattern of early vs. late rising is also more typical of younger speakers. Older speakers, who tend not to use uptalk, do not show the same distinction between early and late rises.

The overall pattern of response choices did not support this hypothesis – participants were no more likely to distinguish between a Question response for early rises and a Statement response for late rises when the vowel was NEAR than when it was SQUARE. Nor, in the more detailed analysis of Question response latencies, was the faster response time for utterances with early rises additionally influenced by the realisation of the vowel, whereas we might have expected the presence of a NEAR vowel to result in an increase in the latency difference between responses to early and late rise stimuli.

However, the analysis of the mouse-tracks suggests that the process of coming to the decision that the utterance was a Question was affected by the quality of the vowel. For both AUC and MD measures, the significant interaction of Vowel and Rise for the Question responses shows that after a NEAR vowel there was less attraction away from the Question response (and towards the Statement response) if the rise started early than if it started late. After the SQUARE vowel there was no such difference – early and late rises were equally good (or bad) cues to the utterance being a Question.

6. Conclusions

Previous research has shown that the NEAR vowel realisation of SQUARE vowels in NZE indexes younger speakers. In the context of 'younger speaker' signalled by a NEAR vowel in the experiment reported in this paper, an early rise onset is more confidently interpreted as indicating a question than a late rise onset. This confidence is reflected in the AUC and MD measures from the mouse-tracking data. This interaction of vowel quality and rise alignment was not apparent in

participants' response selections nor in their response times. This demonstrates the usefulness of mouse-tracking data in categorization tasks: 'the dynamics of action are part and parcel with the dynamics of perception and cognition' [10].

The study also showed, however, a strong overall bias towards Question responses. This made statistical analysis of the detailed properties of the Statement responses problematic. Future work needs to consider other aspects of the rising patterns that were largely ignored in the current study. That is, the simple late rises may not have been sufficient to indicate 'uptalk' (though see Figure 1).

There is also potential to explore more detail of the mouse-track data than has been considered for this paper. For instance, the changes in the trajectory and in the speed of mouse movement during the course of arriving at a decision could be aligned with the points in time at which vowel identity and intonational change information become available.

7. References

- [1] Cruttenden, A., "Falls and rises: meanings and universals", *Journal of Linguistics*, 17:77-91, 1981.
- [2] Grabe, E., "Variation adds to prosodic typology", in B.B.I. Marlien [Ed], *Proceedings of the Speech Prosody Conference*, 127-132, Laboratoire Parole et Langage, 2002.
- [3] Haan, J., *Speaking of Questions: An Exploration of Dutch Question Intonation*. University of Nijmegen, 2002.
- [4] Warren, P., Uptalk - the forms, functions and distribution of a rising intonation, in prep.
- [5] Warren, P. and Britain, D., "Intonation and prosody in New Zealand English", in A. Bell and Kuiper, K. [Eds], *New Zealand English*, 146-172, Victoria University Press., 2000.
- [6] Warren, P., "Patterns of late rising in New Zealand English: Intonation variation or intonational change?", *Language Variation and Change*, 17:209-230, 2005.
- [7] Zwart, J. and Warren, P., "This is a statement? Lateness of rise as a factor in listener interpretation of HRTs", *Wellington Working Papers in Linguistics*, 15:51-62, 2003.
- [8] Hay, J., Warren, P., and Drager, K., "Factors influencing speech perception in the context of a merger-in-progress", *Journal of Phonetics*, 34(4):458-484, 2006.
- [9] Gordon, E. and Maclagan, M., "Capturing a sound change: A real time study over 15 years of the NEAR/SQUARE diphthong merger in New Zealand English", *Australian Journal of Linguistics*, 21(2):215-238, 2001.
- [10] Freeman, J.B. and Ambady, N., "MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method", *Behavior Research Methods*, 42(1):226-241, 2010.
- [11] Schneider, W., Eschman, A., and Zuccolotto, A., *E-Prime Reference Guide*, Psychology Software Tools, Inc., 2001.
- [12] R Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing [<http://www.R-project.org/>], 2014.