



Phonetic differences between uptalk and question rises in two Antipodean English varieties

Paul Warren¹, Janet Fletcher²

¹ Victoria University of Wellington, New Zealand

² University of Melbourne, Australia

paul.warren@vuw.ac.nz, janetf@unimelb.edu.au

Abstract

Analysis of map task data for two varieties of English in which uptalk has long been documented (Australian and New Zealand) indicates differences in the phonetic forms of uptalk rises and question rises. While the details of the phonetic differences are not the same in the two samples, the end result is a more dynamic rise in uptalk than in questions. This difference in rise shape may be indicative of a change-in-progress in the intonational systems of the two varieties.

Index Terms: uptalk, pitch accent height, rise alignment, pitch dynamics, speech roles, varieties of English

1. Introduction

Uptalk (or the use of high rising terminals) is frequently described, particularly in the popular press, as the use of question intonation on statements. It is likely that this view contributes to the common (mis)perception that uptalkers use this type of intonation because they are uncertain of what they are saying. It is a view that is based on two challengeable assumptions. One is that there is a close link between rising intonation and questions. The other, which will be the main focus of this paper, is that the rising patterns found on uptalk are not different from those found on questions.

The first assumption ignores the twin facts that questions frequently do not have intonational rises and that statements can have rising intonation, even in the absence of uptalk. As an example of the first issue, it has been noted that across the world's languages, while intonational marking for yes/no questions is more likely than the use of specific word orders, it is far less common than the use of a question particle [1]. In English, yes/no questions have been shown to more frequently have falling than rising intonation [2], and a negative correlation is reported between the number of morphosyntactic markers of interrogativity and the likelihood of rising intonation [3]. On the second issue, while some varieties of English (e.g. Urban Northern British) routinely have rising intonation on a range of statement types, it is more generally found on statements in most varieties to indicate functions such as listing and continuation, which differ from the typical interactive meanings and functions of uptalk (for an overview see [1]).

With regard to the relationship between uptalk and the types of rising intonation found with questions, we find that expert opinions range from those that see uptalk and question rises to be effectively identical [4-6], through those that note that the forms may be similar [7-9], to those who observe clear

differences. Where differences have been noted, they include aspects both of pitch level or range and of the alignment of rises. In particular, it has recently been noted that some Australian English speakers distinguish statement and question rises through the use of higher pitch accents, i.e. higher starting points for the rise, on questions than on statements [10], though one study reports less consistency in this respect for male than for female speakers [7]. A recent study of Southern Californian English notes that the size of the rise is also linked to discourse function, with rises on floor-holding statements having twice the pitch range of those on non floor-holding statements [11]. Differences in rise alignment have been reported for New Zealand English [12], South African English [13], and Southern Californian [11] (see also [14]). The general finding is that question rises start earlier, relative to the accented syllable, than statement rises. Similar alignment differences have also been reported for Australian English [7, 15], though these are generally not as marked as the differences in the starting pitch level [10, 16, 17].

Forced-choice perceptual studies of Australian and New Zealand English show sensitivity to the differences in pitch accent heights in the former [16] and in rise start alignment in the latter [18]. Australian listeners were most confident of 'statement' judgements when the rise had a low pitch accent and of 'question' responses when it had a high pitch accent. New Zealanders judged early rises in their variety to be more compatible with questions and late rises with statements. Although different means may be used to distinguish question and statement rises in different varieties of English, there may be some commonality to be found in the notion that an uptalk rise is dramatic or dynamic, in comparison with a question rise.

In the current paper, we consider more closely speech data from these two varieties, Australian English (AusE) and New Zealand English (NZE), in a direct comparison of some of the phonetic features of their question and statement rises. In particular we seek to determine whether there is evidence in the phonetic analysis of rises in these varieties to support the previous observations that statement and question rises are distinguished in both varieties, and that this distinction is achieved primarily through pitch-height features in AusE and through the temporal alignment of the rise in NZE.

2. Method

To facilitate our comparison of the realization of rises in the two varieties, we selected speech data from existing corpora employing a common task. This is the map task [19], in which one participant, the Leader, has to explain a marked route to the other, the Follower, so that the latter can reproduce the

route on their own map. Each pair of participants in our datasets completes two map tasks, with pair members changing roles for the second task. The maps seen by Leader and Follower are similar, but have several features that are distinct or that are differently named. This results in interactions that are characterized by both directives and questions. The maps used in the tasks reported below are the same for both English varieties, heightening the similarity of the tasks. An example is given as Figure 10.1 in [1].

2.1. Speech samples

The AusE map task recordings analysed for this paper are a subset of those discussed in [20]. The NZE recordings form part of the New Zealand Spoken English Database (see [21]). The two samples date from around the same time. For the initial analysis reported here, total samples of at least 40 minutes were taken from each corpus. The AusE sample involved two pairs of particularly loquacious speakers each carrying out two map tasks. Twice as many interactions were required from the NZE corpus to get the same coverage. The AusE interactions totaled just under 43 minutes and the NZE interactions around 54 minutes. Speaker sex was balanced for each of the samples.

2.2. Mark-up

The speech recordings for both samples had previously been orthographically transcribed, and the AusE transcriptions had been time-aligned in *xwaves*. Time alignment of the NZE recordings with the orthographic transcriptions was achieved using the web-based implementation of *MAUS*, selecting New Zealand English as the analysis language [22]. Some hand correction of the automatic alignment was necessary.

Databases for the samples from each variety were then produced using *emuR* [23], and consisted of speech files (wav format) with accompanying pitch tracks (generated using the Schaefer-Vincent algorithm in the *wrassp* library [24]) and two text tiers containing time-aligned orthographic word-level transcriptions and pitch events. For the current study, we were interested in the realization of high-rise contours. (Note that the fall-rise contours identified by [25] were not included in the current analysis, which is therefore a conservative study of uptalk in these varieties.) The onset and peak of each high rise had been marked earlier in the *xwaves* analysis of the AusE samples [20], and were transferred to the pitch event tier. The equivalent points in the NZE database had been determined for the analysis presented in [12], and were marked up in the *emuR* database for the current study, using the *EMU-webApp* [26]. Each rise onset and peak was also marked as being on a statement or question, this being determined from a combination of morphosyntactic evidence and the context of the interaction.

For each rise in the two databases, two additional points were marked in *EMU-webApp*. These were the onset and end of the nuclear accent unit. An additional feature was noted during the labelling process, which was the tendency, primarily in the NZE data, for a rise to consist of a step-up in pitch from the accented syllable to a following unaccented syllable, e.g. from the first to the second syllable in *castle*. In these cases, the low point for calculating the size of the rise was taken to be the end of the first syllable, but the alignment point for the rise was taken to be the beginning of the second syllable, since this is the point at which the hearer becomes aware of the rise in pitch. This affected 10.2% of NZE rises

and 1.4% of AusE rises, and was found for both statements (6.0%) and questions (4.8%).

2.3. Data

The overall counts of high rises in the samples are given in Table 1. More high rises are produced by speakers in the Leader role than in the Follower role, but it should be noted that Leaders generally speak for a much higher proportion of the recording time than Followers. In addition, the relative proportion of statement rises to question rises is higher for Leaders (who are issuing directives) than for Followers (who often check for information, using questions).

Table 1. *Counts (with proportions) of high rises.*

Role	Type	AusE	NZE
Leader	Statement	198 (0.80)	136 (0.64)
	Question	48 (0.20)	76 (0.36)
Follower	Statement	68 (0.59)	32 (0.35)
	Question	48 (0.41)	59 (0.65)

A number of measures were taken from the marked-up files in *emuR*. Pitch measures included pitch height at the rise onset and peak. Pitch values (Hz) were transformed to the ERB scale for the analysis below, since this scale provides some normalization for individual pitch range differences. Time measures included the onset and offset of the accent unit, and the onset and peak of the rise. Derived measures from these values included pitch range of the rise, pitch rate-of-change across the rise, accent unit duration, rise duration, and the position of the rise onset as a proportion of the duration of the accent unit.

2.4. Statistical analysis

For the statistical analysis of the results presented below, linear mixed effects models were built in R using the *lme4* package [27], and the statistical significance of including a factor or interaction in a model was assessed using the *mixed* command from the *afex* package [28]. In each model the dependent variable (DV) was one of the measures above. The predictors or independent variables (IVs) were Variety (AusE or NZE), Role (Leader or Follower), Sex (female or male speaker), and Sentence Type (Statement or Question), and their interactions. Speaker was included as a random effect. Because there was considerable multicollinearity between the predictors, all IVs were centered before models were built. Some of the collinearity is not surprising – for instance, we would expect Role to be correlated with Sentence Type because Leaders tend to give instructions using statements and Followers tend to seek confirmation or further information using questions.

3. Results

3.1. Pitch height

Three measures of pitch height (in ERB units) are considered: the height of the onset of the rise, the peak height at the end of the rise, and the size of the rise (i.e. the difference between the first two measures).

The statistical analysis for the pitch height at the onset of the rise revealed significant simple effects for Sex ($\chi^2(1)=50.19$, $p < 0.0001$), Role ($\chi^2(1)=3.72$, $p = 0.05$) and

Sentence Type ($\chi^2(1)=17.35$, $p < 0.0001$), and a significant interaction of Role and Sex ($\chi^2(1)=5.24$, $p < 0.05$). This interaction is shown in Figure 1, where the simple effect of Sex is also clear. The interaction effect results from the presence of a Role effect for male speakers but not for females. When males are in the Leader role they have higher starting points for their rises than when they are in the Follower role. Other research has noted heightened prosodic marking for speakers issuing rather than following instructions, in a game-based task [29].

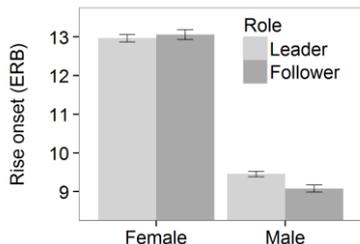


Figure 1: Rise onset: interaction of Sex and Role.

The Sentence Type effect reflects a higher starting point for questions than for statements. The difference is small but highly significant, and does not interact with Variety. In terms of our initial research question, we observe that in both AusE and NZE there is a difference in the onset pitch level in statements and questions.

The analysis of the final pitch value of the rises indicates an overall higher peak for questions than for statements ($\chi^2(1)=8.75$, $p < 0.005$). There were also significant simple effects for Role ($\chi^2(1)=31.08$, $p < 0.0001$) and Sex ($\chi^2(1)=45.13$, $p < 0.0001$) and significant interactions of Variety with Sex ($\chi^2(1)=4.51$, $p < 0.05$) and with Role ($\chi^2(1)=6.98$, $p < 0.01$) and a three-way interaction of Role with Variety and Sentence Type ($\chi^2(1)=10.24$, $p < 0.001$). This three-way interaction is shown in Figure 2. What is clear from this figure is that the general tendency for question rises to have higher endpoints than statement rises is not found for AusE speakers in the Leader role, where their statement rises finish at a considerably higher level than their question rises.

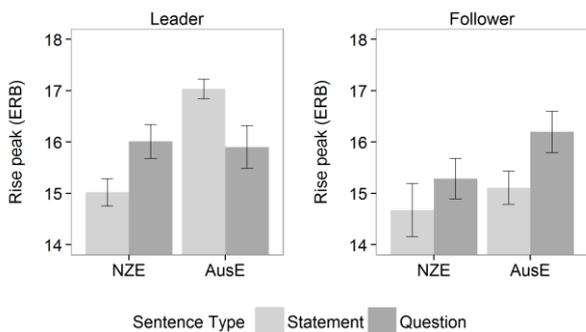


Figure 2: Rise peak: interaction of Role, Variety and Sentence Type.

The interaction of Variety and Role reflects the fact that AusE speakers have much larger peak values for their rises as Leaders than as Followers (largely due to the effect noted above for Leader statements), while NZE speakers have little difference between Leader and Follower peaks. The

interaction of Variety with Sex arises because while women show little difference between the two samples (the rise peaks for AusE women are some 0.2 ERB higher than those for NZE women), the men exhibit a larger difference in the same direction (1.7 ERB). Given the small number of speakers in the samples, this difference might be due to individual characteristics rather than more general varietal differences. Our planned analysis of a larger set of speakers will address this possibility.

Looking now at the overall size of the rises we find a significant simple effect of Role ($\chi^2(1)=17.20$, $p < 0.0001$), with larger rises for Leaders than for Followers. There was also a three-way interaction of Role, Variety and Sentence Type ($\chi^2(1)=9.20$, $p < 0.005$; see Figure 3). While questions generally have larger rises than statements, this pattern is reversed for AusE speakers in the Leader role, as we saw also for the rise peak values. Note also that the differences between question and statement rises are much reduced for the NZE speakers compared to those for the AusE speakers.

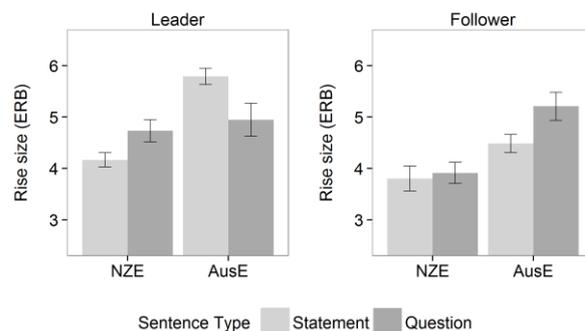


Figure 3: Rise size: interaction of Role, Variety and Sentence Type

3.2. Temporal measures

Two temporal measures are reported: the alignment of the starting point of the rise, and the duration of the rise.

Rise alignment was calculated as a proportional measure determined by where the rise started relative to the entire duration of the accent unit, with the latter defined as the period from the beginning of the nuclear accented syllable to the end of the intonation group. Proportional data were used to normalize for differences in the durations of accent units. The analysis returned a four-way interaction of Sentence Type, Variety, Role and Sex ($\chi^2(1)=5.09$, $p < 0.05$). This four-way interaction is shown in Figure 4. It can be seen from this figure that the alignment point is generally later for statements than for questions (as reflected in a significant simple effect for Sentence Type: $\chi^2(1)=69.24$, $p < 0.0001$), with the exception of the Australian males when in the role of Follower, where there is a small difference in the opposite direction. The size of the delay for statement rises relative to question rises is variable, and is on the whole more marked for the NZE speakers than for the AusE speakers, the exception being the NZE females in the Follower role, where the difference is in the same direction but less distinct. (Note that this set includes only 13 statements and 30 questions.)

A significant interaction of Sentence Type and Variety ($\chi^2(1)=9.37$, $p < 0.005$) reflects the larger difference for NZE speakers (statements: 0.51, questions: 0.32) than for AusE speakers (statements: 0.51, questions: 0.42). An interaction of

Sentence Type with Role ($\chi^2(1)= 6.93, p < 0.01$) arises because Leaders (statements: 0.52, questions: 0.33) make the distinction more clearly than Followers (statements: 0.47, questions: 0.40).

It is also evident from Figure 4 and from the interaction of Sentence Type and Variety that the distinction between the AusE and NZE data lies in the earlier rise starts for the NZE questions, whereas the two varieties have comparable rise starts for statements. An apparent-time pattern of earlier rises for questions from younger speakers relative to older speakers was noted in earlier research on NZE [12].

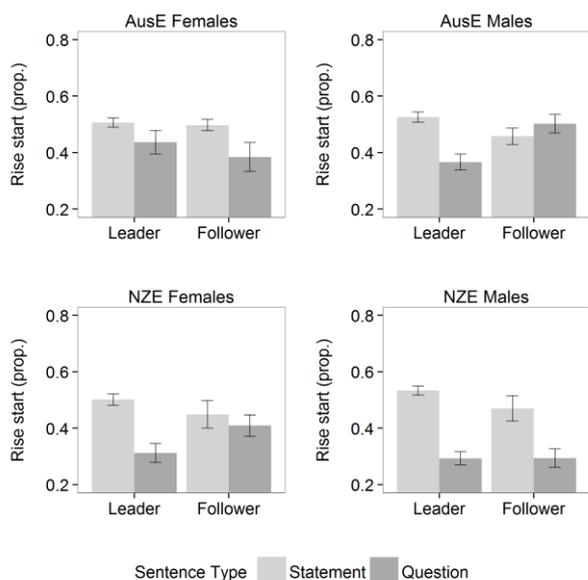


Figure 4: *Rise start alignment: interaction of Variety, Sex, Role, and Sentence Type*

Analysis of rise durations (again normalized for accent unit duration) shows a significant effect for Sentence Type (question rises > statement rises: $\chi^2(1)= 83.75, p < 0.0001$) and a significant interaction of Sentence Type and Variety. While both varieties have longer rises for questions, the difference is greater for NZE (0.45 vs. 0.28) than for AusE (0.40 vs. 0.31).

3.3. Rate-of-change

Taken together, the two sets of differences between statement and question rises, namely pitch differences (stronger for our AusE samples) and timing differences (stronger for our NZE samples), indicate a more dramatic rise for statements than for questions. Indeed, an analysis of rate-of-change in pitch across the rise (ERB/normalized duration) showed significant effects of Sentence Type (statement rises have a greater rate-of-change than question rises: $\chi^2(1)= 30.30, p < 0.0001$) and of Role (Leaders have higher values than Followers: $\chi^2(1)= 11.02, p < 0.001$). Importantly, Variety does not interact with these factors. In other words, speakers of both varieties use more dynamic rises in statements than in questions, though the means of achieving this differs between the speaker groups sampled for this study.

4. Discussion

The analysis above has provided partial support for previous observations on how AusE and NZE distinguish uptalk and

questions. Earlier studies indicated that AusE uses pitch distinctions rather than differences in the temporal alignment of the rise, while the opposite has been claimed to be true of NZE. The analysis of the pitch data above revealed stronger pitch effects for AusE than for NZE. The specific observation that the rise follows a lower pitch accent for statements than for questions in AusE was not supported, but it should be noted that the analysis to date has only included simple rises, and not the fall-rise patterns (and in particular compound or split fall-rises) that have recently been linked to uptalk in AusE. Nevertheless, both the pitch peak and the size of the rise showed clear effects, but effects that depended also on the role of the participant in the map task. In particular, Leaders showed precisely the kind of pattern observed in earlier studies, namely greater pitch marking of the rise in statements than in questions. The opposite pattern was found for AusE Followers, but note that the rise and peak values for Followers' questions are still less extreme than those of the Leaders' statements. This is likely to be related to the nature of the Leader's role. Previous discourse analyses of the same dialogues has shown that the Leaders produce more initiating moves and forward-looking dialog acts than Followers [20], and so are using a wider pitch range in these initiating moves.

The analysis of the temporal features of the rise is similarly compatible with previous research, particularly on NZE. The earlier rises for questions have been noted in production studies and also in perceptual experiments. These earlier rises are a feature of the speech of younger NZE speakers, and are more likely to be interpreted as questions when other speech features indicate that the speaker is likely to be a younger speaker [30]. A recent study of rise perception in AusE as spoken in Melbourne showed that teenage speakers are also likely to hear more questions when the rise is both earlier and larger [15].

We conjecture that the differences between the two varieties are in how speakers choose to realise the more dramatic rises (as reflected in rate-of-change values) associated with uptalk. The recent perceptual studies cited above suggest that listeners may be sensitive to the higher rate-of-change values found in each variety.

5. Conclusions

Uptalk has been defined as “[a] manner of speaking in which declarative sentences are uttered with rising intonation at the end, as if they were questions” [31]. There is growing evidence, to which the analyses in the current paper contributes, that speakers in fact have different rise patterns for uptalk and questions. While the differences are often fairly subtle in the data analysed here, they are nevertheless statistically significant. Often, of course, the identification of a rise as marking a question or a statement will be obvious from the utterance context [6]. In varieties with a higher incidence of uptalk, however, there may be increasing pressure to make a realisational distinction [32], which may in turn evolve into a phonological contrast. Both AusE and NZE have a fairly high incidence of uptalk (19% in one report [25]), and so the phonetic differences documented in this paper may be a reaction to such pressure. Further analysis of different age groups in the map task corpora for both AusE and NZE (complementing analyses presented for NZE in [12]), as well as of more recent recordings of younger speakers, will help us to track this potential change-in-progress in the intonational systems of these two varieties.

6. References

- [1] P. Warren, *Uptalk: The Phenomenon of Rising Intonation*. Cambridge: Cambridge University Press, 2016.
- [2] C. Bartels, *The intonation of English statements and questions: a compositional interpretation*. New York: Garland Publishing, 1999.
- [3] E. Grabe, "Variation adds to prosodic typology," in B. Bel and I. Marlien (eds.), *Proceedings of Speech Prosody 2002*, 127-132, Aix-en-Provence: Laboratoire Parole et Langage, 2002.
- [4] D.R. Ladd, *Intonational Phonology*. Cambridge: Cambridge University Press, 1996.
- [5] R. Lakoff, "Language and woman's place," *Language in Society* 2:45-79, 1973.
- [6] G. Guy, et al., "An Intonational Change in Progress in Australian English," *Language in Society*, 15(1):23-51, 1986.
- [7] J. McGregor, *High Rising Tunes in Australian English*. PhD thesis, Macquarie University, 2005
- [8] P. Bryant, *Australian Questioning Intonation: An Addition to Speakers' Response-Seeking Repertoire*. BA(Honours) thesis, Australian National University, 1980
- [9] D. Britain, "Linguistic change in intonation: The use of high rising terminals in New Zealand English," *Language Variation and Change*, 4:77-104, 1992.
- [10] J. Fletcher and J. Harrington, "High-Rising Terminals and Fall-Rise Tunes in Australian English," *Phonetica*, 58(4):215-229, 2001.
- [11] A. Ritchart and A. Arvaniti, "The form and use of uptalk in Southern Californian English," *Proceedings of Meetings on Acoustics*, 20, 2014.
- [12] P. Warren, "Patterns of late rising in New Zealand English: Intonation variation or intonational change?," *Language Variation and Change*, 17:209-230, 2005.
- [13] N. Dorrington, *'Speaking up': A Comparative Investigation into the Onset of Uptalk in General South African English*. BA(Honours) thesis, Rhodes University, 2010
- [14] M.E. Armstrong, P. Piccinini, E., and A. Ritchart, "The phonetics and distribution of non-question rises in two varieties of American English," in M. Wolters, et al. (eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*, Paper number 0927. 1-5, Glasgow, UK: University of Glasgow, 2015.
- [15] C. Holt and J. Fletcher, "Perception and interpretation of low onset rising tunes by prelingually deaf cochlear implant users," in M. Wolters, et al. (eds.), *Proceedings of the 18th International Congress of Phonetic Sciences*, Paper number 0188. 1-5, Glasgow, UK: University of Glasgow, 2015.
- [16] J. Fletcher and D. Loakes. *Interpreting rising intonation in Australian English*. in *Proceedings of Speech Prosody 2010*. University of Illinois Press, 2010.
- [17] K. Webb, *High Rising Terminals in Australian English: Form and Function*. BA(Honours) thesis, University of Melbourne, 2008
- [18] P. Warren, "Sociophonetic and prosodic influences on judgements of sentence type," in J. Hay and E. Parnell (eds.), *Proceedings of the 15th Australasian International Conference on Speech Science and Technology*, 185-188, Christchurch: ASSTA, 2014.
- [19] A. Anderson, et al., "The HCRC Map Task Corpus," *Language and Speech*, 34:351-366, 1991.
- [20] J. Fletcher, et al., "Intonational Rises and Dialog Acts in the Australian English Map Task," *Language and Speech*, 45(3):229-253, 2002.
- [21] P. Warren, "NZSED: building and using a speech database for New Zealand English," *New Zealand English Journal*, 16:53-58, 2002.
- [22] F. Schiel, "Automatic Phonetic Transcription of Non-Prompted Speech," in J.J. Ohala, et al. (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences*, 607-610, San Francisco: Regents of the University of California, Berkeley, 1999.
- [23] R. Winkelmann, et al., *emuR* vsn.0.0.10. Available from: <https://github.com/IPS-LMU/emuR>, 2015.
- [24] R. Winkelmann, L. Bombien, and M. Scheffers, *wrassp* vsn.0.1.3. Available from: <https://github.com/IPS-LMU/wrassp>, 2015.
- [25] J. Fletcher and D. Loakes, "Patterns of rising and falling in Australian English," in P. Warren and C.I. Watson (eds.), *Proceedings of the Eleventh Australasian Conference on Speech Science and Technology*, 42-47, Canberra: Australasian Speech Science and Technology Association Inc., 2006.
- [26] R. Winkelmann, *EMU-webApp* vsn.0.0.28. Available from: <https://github.com/IPS-LMU/EMU-webApp/>, 2015.
- [27] D. Bates, et al., *lme4* vsn.1.1-8. Available from: <http://CRAN.R-project.org/package=lme4>, 2015.
- [28] H. Singmann, B. Bolker, and J. Westfall, *afex: Analysis of Factorial Experiments* vsn.0.13-145. Available from: <https://cran.r-project.org/web/packages/afex/index.html>, 2015.
- [29] S.R. Speer, P. Warren, and A.J. Schafer, "Situationally independent prosodic phrasing," *Laboratory Phonology*, 2(1):35-98, 2011.
- [30] P. Warren, "The interpretation of prosodic variability in the context of accompanying sociophonetic cues," under review.
- [31] Oxford Dictionaries. 2015 11 Nov 2015; Available from: <http://www.oxforddictionaries.com/definition/english/uptalk>.
- [32] J. House, "Constructing a context with intonation," *Journal of Pragmatics*, 38(10):1542-1558, 2006.