Intonational variation in adolescent conversational speech: rural versus urban patterns

Janet Fletcher and Deborah Loakes

Department of Linguistics and Applied Linguistics
University of Melbourne, Australia
{janetf; dloakes}@unimelb.edu.au

Abstract

The conversational speech of ten female adolescents was analyzed with a view to determining whether there is intonational variation between rural and urban varieties in Australian English. The data revealed that urban females use marginally more 'uptalk' (i.e. high rising terminals) than their rural counterparts with syntactic declarative utterances. Aspects of intonational variation are presented in terms of a prevailing intonational model of English, and discourse annotation schema.

1. Introduction

Recent work on English intonation suggests that variation between and within varieties is prevalent, particularly in countries like the United Kingdom. These differences have been well-documented in recent years and include semantic, systemic, and realization differences in pitch accent category and boundary type. For example, Grabe et al. [1] found a high level of variation in nuclear tunes between southern varieties of British English and northern varieties, echoing earlier findings by Cruttenden [2]. Whilst falling intonation is the dominant declarative statement melody in southern varieties of British English, the reverse is the case in Belfast English, where statement intonation is predominantly rising. This is precisely the kind of dialectal intonational variation that research on speech technology applications like text-to-speech and spoken dialog systems needs to take into account.

Like Belfast English, Australian English has also been referred to as a rising variety ([3]). "Uptalk" or the use of rising and high pitch at the end of syntactic declarative statements is relatively common in certain types of interactive discourse in Australian English. In a number of recent corpus-based prosodic studies ([4,5]), it was found that speakers tend to use a variety of 'non-questioning' high rises include complex rises (e.g. expanded range fall-rises or rise-fall-rises) as well as simple terminal rises. It was also found that many of the terminal high (and low) rises used by speakers are generally part of a split or compound fall-rise, i.e. the terminal rise is the final part of a divided nuclear tone that is realized across the entire intonational phrase.

These findings were based on extensive analyses of map task corpora. The map task is a cooperative and collaborative task and it is perhaps not surprising that there is greater use of rising versus falling intonation in map task interactions. In other words, uptalk is more indicative of a 'continuative' rather than questioning function, and is associated with the 'encouraging' semantic nuances generally attached to fall-rises in other varieties of English. More generally, Warren [6] cites work by Meyerhoff [7] who suggests that uptalk establishes 'common ground' between speakers and hearers in an interaction or signals 'group membership'. It has also been suggested that uptalk is a kind of checking device to ensure that the other participant(s) in an interaction are in tune with the 'mutual construction of a conversation' ([6] p. 210).

Given the context of the map task as a collaborative interaction, it is perhaps not surprising that 59% of all terminal tunes in one of the earlier corpus studies of Australian English were high rising [5]. One of the key findings was that there appears to be no significant difference between males and females with respect to the incidence of uptalk, although there is a high level of variation amongst different speakers, irrespective of speaker sex. The corpus consisted of speakers general Australian English from the Sydney area in New South Wales, who were aged 35 and over. These earlier results were surprising in view of previous sociolinguistic work from the '80s [3] which suggested that uptalk is generally used by adolescent females. If young women are meant to be the heralders of linguistic change (e.g. [8]), then this suggests that uptalk may now be a feature of both male and female intonational patterns, unlike twenty years ago. In a study of a closely-related variety, Warren [6] reports that the phenomenon is still more widespread in younger versus older speakers of New Zealand English. In a recent study of rising intonation in the speech of Sydney adolescents, McGregor [9] found that rising intonation is as prevalent in both female and male interactions. Like Warren [6], and previous studies by the first author of this paper [4, 5], McGregor's study was based on map task interactions.

The aims of the current investigation were to extend previous studies and look at adolescent discourse in metropolitan Melbourne and in rural Victoria, in Australia. We also chose to focus on free conversation between groups of young adolescents who know each other well. Our goal was to see whether uptalk is as prevalent in conversational discourse between young adolescents who are friends, as in constructed collaborative tasks like the map task. We also wanted to examine general tune usage across the corpus, rather than focusing solely on high rises. It is important to test the frequency of uptalk relative to falling tunes, for example, to get a sense of the overall system of tune usage in a particular variety and in a particular type of speech situation. Our other major goal was also to examine Australian English as spoken in Melbourne and Victoria, to get a sense of any potential regional variation that has anecdotally been reported elsewhere. It has often been claimed that there is relatively little regional variation in Australian English (e.g.[10]) although there is an urban/regional division that is apparent in vowel formant patterns (e.g. [11]). No experimental research has been undertaken to examine evidence for this split in the intonational and prosodic characteristics of urban versus regional speech patterns in Australia.

2. Methods and materials

2.1. Corpus

Recordings from an existing corpus of English spoken in Victoria, Australia were analyzed in this study. The corpus was recorded as part of the 'Dimensions of Australian English' project at Monash University, Melbourne. Two sets of conversations between five adolescent girls were selected. One set of girls were all school friends and resided in Tarrington, Victoria, a rural community to the West of Melbourne that had strong patterns of German settlement in the nineteenth century. The urban group consisted of five girls living in Mount Waverley in the eastern suburbs of Melbourne. For both groups, an entire fifteen minute conversation was used for analysis in this study.

2.2. Word and Prosodic Labelling

The original recordings were digitized at 22 KHz., and the acoustic waveform files and F0 signal were annotated according to ToBI (Tones and Break Indices) conventions that have been adapted for Australian English [4]. Word boundaries were identified and orthographically annotated. Major pitch movements corresponding to pitch accents and intermediate and intonational phrase boundaries were labelled using the F0 signal and auditory analysis. The specific tone sequences under investigation are summarized in Table 1. The tunes listed in column 2 of Table 1 are those that are generally classified as uptalk [4]. All have high phrase accents and high boundary tones. Intermediate phrases that terminate with final H- tones, that are not intonational phrase-final, can also be classified as uptalk (e.g. H* H- tone sequences).

Table 1 Summary of tunes and associated ToBI annotation categories

	High Rise (Uptalk)	***Low Rise***	***Mid-level***	***Fall***
Simple (low pitch onset - nucleus)	*L* H-H%* *L* H-*	*L*L-H%*		
(high pitch onset - nucleus)	*(L*)H*H-H%* *H* H-*		*H* H-L%* *H* H-*	*H* L-L%* *H* L-*
Complex(fall-rise)	*H*+L H-H%*	*H*L-H%*		
compound	*H* L*/!H% H-H%*	*H*L*/!H* L-H%*		

Figure 1 shows an example of a syntactic declarative utterance "we usually get jobs to do on Saturdays" realized with a high rising terminal or uptalk. Each tune was also labelled using broad talk-in-interaction criteria, namely position in turn. Tunes were labelled as either turn-internal (i.e. floor holding) or turn-yielding.

2.3 Micro-level discourse coding: dialog acts

The coding for dialog acts used in this study was based on the modified DRI/DAMSL scheme used in an earlier study [5]. The dialog act coding system adopted in that study, SWBD-DAMSL was used because it permits a relatively fine-grained analysis of different dialog acts beyond the informal semantic categorization of the high rising tune that has been carried out in earlier studies of Australian English [4]. For example, *no* answers are coded separately from *yes* answers.

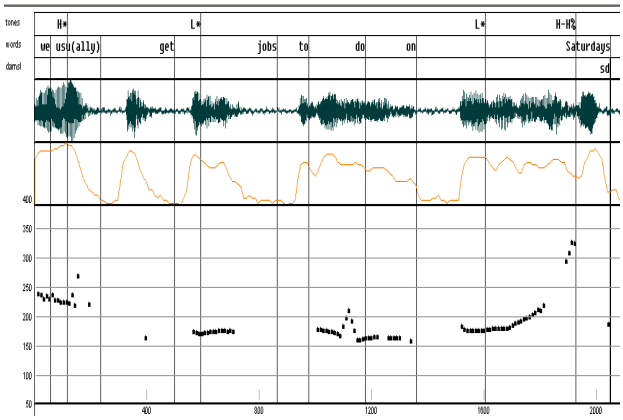

Figure 1. An example of a declarative utterance realized as a high-rising tune with a low pitch onset (L* H-H%).

Table 2 summarizes some of the main 'forward' and 'backward' communicative functions used in the dialog act coding of the conversations. Tunes were categorized according to dialog act, and Chi-square analyses performed to compare interaction between tune type and turn position, on the one hand, and tune type and dialog act, on the other.

Table 2. Some of the major SWBD/DAMSL codes for forward and backward communicative functions

Forward-communicative-functions	**Backward-communicative-functions**
Statement *sd* *sv* *Information requests* *qy* *qw* *ad*	*Agreement* *a* *ar* *Understanding* *br* *b* *bk* *Answer* *ny* *nn*

3. Results

3.1. Tune distribution and position in turn

Table 3 summarizes the distribution of the main tunes in the corpus for the urban and rural speakers. By far the most prevalent intonational phrase-final tune in the urban girls' interactions was a mid-level tune (H* H-L%), with 34% of intonational phrases (279) terminating with this intonational pattern. The next most common tune was a final fall (H* L-L%) at 22%. High rising tunes (irrespective of whether they were yes/no question rises or uptalk) accounted for only 23% of all intonational phrases, with the rest consisting of fall-rise or low rising tunes.

In the rural speakers' corpus, 47% of all intonational phrases (217) terminated in final falls (H* L-L%), and 26% terminated in high rises. Unlike the urban corpus only 9% of tunes were mid-level (H* H-L%), whereas 17% of tunes were fall-rises.

Table 3. Distribution of major tunes across city and rural corpus expressed as proportion of total

Tune	Urban n=279	Rural n=212
H* H-H%	8%	18%
L* H-H%	10%	7%
H* L* H-H%	5%	1%
L* L-H%	15%	1%
H*L-H%	6%	17%
H* H-L%	34%	9%
H* L-L%	22%	47%

Figures 4 and 5 show tune distribution, according to position in the conversational interaction. The rural corpus (Figure 5) consisted of fewer turns overall, with 73 floor-holding and 44 turn-yielding tunes. There was overall all less talk in the rural girls' interactions, compared to the urban corpus. The interaction between the type of tune and turn function approached significance (Chi-sq=9.92; $p<0.07$) in the rural corpus. The urban corpus, by contrast consisted of more than twice as many turns, and more talk overall. 165 intonational phrases were turn internal (i.e. floor-holding) and 94 were turn-yielding, but there was no significant interaction between turn position and tune type ($p>0.05$). However, if we examine the tune distribution more closely in the urban corpus (Figure 4), relatively more rising tunes concluded intonational phrases that were turn-internal, than turn-final. This was markedly so for mid-level, rise-fall, and low rise tunes. These tunes have often been described as performing continuative functions in spoken English discourse, so this result is not particularly surprising. Amongst the high rising tunes, there was very little difference in distribution according to turn position. The same was true for falling tunes.

In the rural corpus, rising tunes (i.e. simple rises, complex rises, and mid-level tunes) were also more likely to be floor-holding, whereas falling tunes were more or less evenly divided between floor-holding and turn-yielding functions. The major difference between the two data sets was the prevalence of mid-level tunes in floor-holding contexts in urban versus rural talk-in-interaction. That is, rural speakers were less likely to use a mid-level tune in floor-holding contexts than urban speakers. These ties in with the overall differences in the distribution of this tune between the two data sets reported above.

3.2 Correspondence between uptalk and dialog act

Figures 6 and 7 show the dialog act/high rise distribution in the urban and rural corpus. In the urban corpus, there was a significant interaction between tune type and dialog act (Chi-sq=39.63; $p<0.0001$). There was a tendency for yes/no questions to terminate with high onset high rises, although these were also realized as low onset high rises in many instances. Relatively fewer non-question dialog acts terminated with high onset high rises, yet more low onset high rises concluded syntactic declarative utterances. In other words, both types of high rise were associated with non-question dialog acts, i.e. were representative of uptalk. High-level or mid-level tunes were far more prevalent than rising

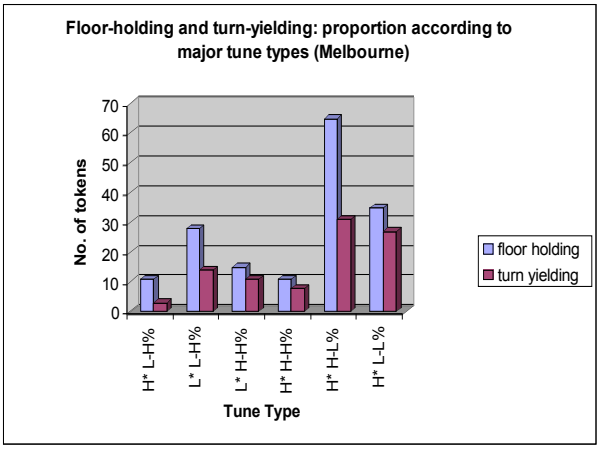

Figure 4. *Distribution of major tunes in urban corpus according to broad discourse function (turn-yielding or floor-holding).*

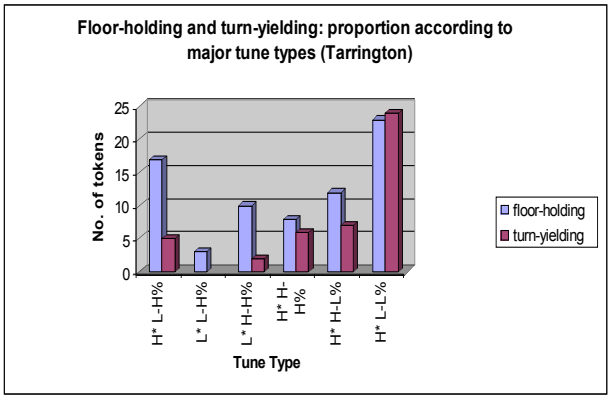

Figure 5. *Distribution of major tunes in rural corpus according to broad discourse function (turn-yielding or floor-holding).*

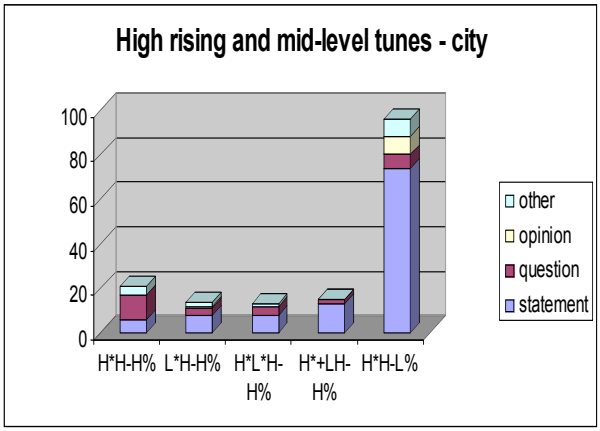

Figure 6. Tune/dialog act correspondence for mid-level and high rising tunes in urban corpus

tunes, by contrast, and a much greater proportion of these tunes concluded statements or opinion type dialog acts.

There was a strong tune/dialog act interaction in the rural corpus (Chi-sq=33.58, $p<0.001$) although there were clear differences in the patterns of distribution. Like the urban girls, rural girls tended to use mid-level tunes to express opinions and statements, whereas the use of high rising tunes with questions was evenly spread between low onset and high onset high rises. Both urban and rural girls used uptalk (i.e. high rising terminals with syntactic declaratives), with urban girls using more rising tunes overall. However, as mentioned earlier, urban girls produced far more talk than rural girls (roughly twice as many turns, and 60 more intonational phrases). Urban girls used more mid-level tunes with statement-like utterances than rural girls. However, recall from the discussion in 3.1 that rural girls were also much more likely to use falling intonation than urban girls. Although not presented in Figure 7, falling tunes corresponded largely with opinion or statement dialog acts, and certain types of interrogatives (usually with question words). A similar pattern of distribution was evident for falling tunes used by urban girls, even though falling intonation was less prevalent than mid-level or rising intonation in their interactions.

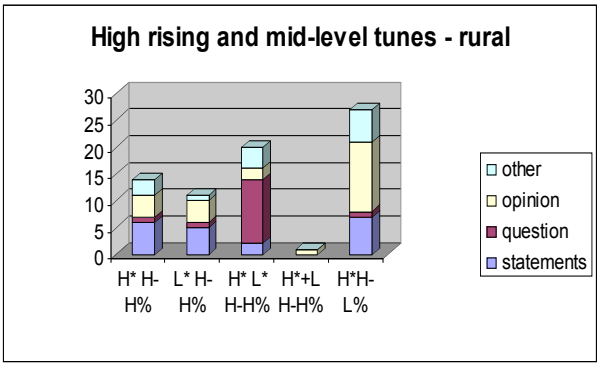

Figure 7. Tune/dialog act interaction for mid-level and high rising tunes in rural corpus

4. Discussion and conclusions

Regional variation is clearly apparent in the range of intonational tunes used by the young urban and rural speakers of Australian English investigated in this preliminary study. More conversational interactions will need to be analyzed to support the conclusions of this study, but it is clear that all speakers used uptalk to a varying degree in the talk-in-interaction observed, although the percentage of rising to falling intonation differed markedly between the two data sets. Of all tunes produced by urban speakers, 88% were rising, although only 23% of tunes could be classified as high rising with most of these tunes constituting uptalk (i.e. high rises with syntactic declarative utterances). Many of these high rises were turn-internal and floor-holding, although the most common floor-holding tune tended to be mid-level. By contrast, nearly half of all intonational tunes in the rural data set were falling. Of the high rising tunes, only two were associated with yes/no questions, and most terminated syntactic declarative utterances. More rises were in floor-holding position than in turn yielding position, although the predominant floor-holding rise tended to be the fall-rise tune, i.e. the traditional "continuation" tune in English, whereas in the urban corpus, the mid-level tune appeared to perform this function.

These results differ markedly from earlier work on uptalk based on map-task interactions [e.g. 4,5], where more than half of all tunes used by speakers constituted uptalk. However, we analyzed spontaneous talk-in-interaction in this study, and the distribution of dialog acts also differed markedly from the earlier studies [5]. Map task interactions tend to consist of many instruction and checking dialog acts and information requests, where as in the spontaneous conversations observed in this study, there were few requests for information, and largely statement and opinion dialog acts. Earlier studies of Australian English intonation have noted that uptalk is more prevalent in semantically difficult tasks, than in normal everyday conversational interaction (see [4] and [6] for summaries of this research). Whilst uptalk is a feature of both the urban and rural girls' intonational patterns, the overall pattern of tune usage suggests a degree of intonational variation, beyond the dominance of the high rising terminal, particularly in spontaneous talk-in-interaction. We intend to analyze more data to confirm these trends.

On a methodological note, the combination of ToBI annotation with DAMSL coding used in this study, confirms that different intonational tunes in a language variety can be used with similar types of utterance or dialog acts, and conversely, the same tune (e.g. a high rising terminal, or mid-level tune) can occur with different utterance types (e.g. statements, opinions, or questions). Furthermore, the ToBI annotation criteria for English intonation, as applied to Australian English, allow the analyst to test tune/function interactions for a large number of rises (see Table 2) and tunes in general. Earlier studies have tended to use traditional contour-based intonational models (see [2] for a summary) that do not allow a fine distinction between low and high onset rises, for example. Our results support a componential approach to intonational meaning [12], whereby pitch accents (low or high) combined with different boundary tones (low or high), contribute in different ways to discourse interpretation.

5. References

[1] Grabe, E. et al.. 2000. Pitch accent realization in four varieties of British English. *Journal of Phonetics* 28:161-195

[2] Cruttenden, A. 1987. *Intonation*. Camb.: CUP

[3] Horvath, B. 1985. *Variation in Australian English*. Camb.:CUP.

[4] Fletcher, J. and Harrington, J. 2001. High rising terminals and fall-rises in Australian English *Phonetica, 58(4),* 215-229.

[5] Fletcher, J. 2005. Compound rises and uptalk. *Interspeech 2005*, 4pp.

[6] Warren, P. 2005. Patterns of late rising in New Zealand English: intonational variation or intonational change? *Language Variation and Change* 17, 209-30.

[7] Meyerhoff, M. 1991. *Grounding and overcoming obstacles; the positive politeness motivations of high rise terminals*. Manuscript. Victoria University Wellington.

[8] Labov, W. 1972. *Language in the inner city*. Philadelphia: University of Pennsylvania Press.

[9] McGregor, J. *High rising tunes in Australian English*. Unpublished PhD thesis, Macquarie University.

[10] Mitchell, A. and Delbridge, A. 1965. *The speech of Australian Adolescents*. Sydney: Angus and Robertson

[11] Cox, F. 1999. Vowel change in Australian English.*Phon* 56, 1-27.

[12] Pierrehumbert, J. and Hirschberg, J. 1990. The meaning of intonation in the interpretation of discourse. In. P.Cohen et al. Eds.) Intentions in communication. Camb.:MIT. 271-311.