DISTRIBUTION AND ALIGNMENT OF F0 CONTOURS IN TAMIL

Elinor Keane

Oxford University Phonetics Laboratory
Elinor.keane@phon.ox.ac.uk

ABSTRACT

Previous work on Tamil intonation suggests that each word in a phrase except the final verb typically bears a fall-rise-fall f_0 contour. The distribution of these contours was investigated in more detail by recording eighteen speakers reading sentences containing nouns of varying length in phrase-medial and phrase-final positions. This established that phrase-final nouns can bear fall-rise-fall contours but are not required to do so, and revealed the possibility of longer words bearing a double fall-rise-fall pattern.

The alignment of the f_0 turning-points was measured to investigate whether the peak is better characterized phonologically as the trailing tone of an L*H accent or a boundary tone. The balance of evidence pointed to the high tone being associated with the boundary of a low-level constituent, maximally the prosodic word.

Keywords: Tamil, intonation, f_0 contour, distribution, alignment.

1. INTRODUCTION

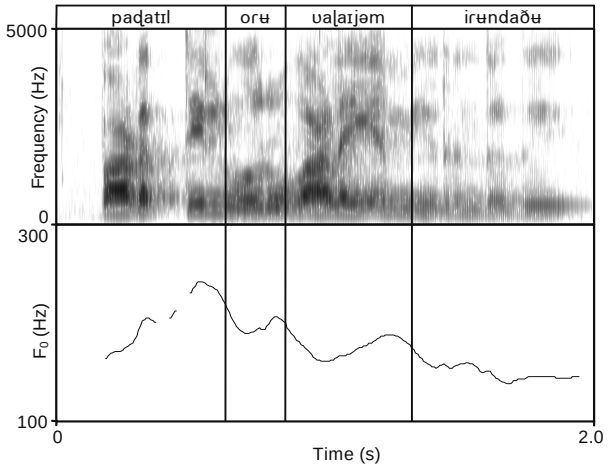
F_0 traces of sentences in Tamil can be described informally as a series of peaks. In short sentences words are typically marked by a peak towards the end of the word, usually preceded by a low f_0 turning-point towards the beginning of the word. This produces a characteristic fall-rise-fall contour, seen on the second and third words in figure 1, a spectrogram and f_0 contour for the sentence:

paɖatɨl oɾʉ ʋaɭaɪjəm iɾʉndaðʉ
picture.loc a bracelet be.past.3sn
'There was a bracelet in the picture.'

As in figure 1, phrase-final words in statements are consistent exceptions to the fall-rise-fall pattern, being marked instead by gradually declining f_0. It is unclear whether this pattern is prosodically determined or a property of the verb that occupies phrase-final position (Tamil word order being SOV).

Another unresolved issue concerns the alignment of the f_0 turning-points in the fall-rise-fall contour, and this has implications for its phonological characterization. Low turning-points have been found consistently within word-initial syllables ([4, 5]). There is no lexically distinctive stress in Tamil, but since word-initial syllables are also marked by distributional asymmetries and weak phonetic prominence they may bear lexical accent. A plausible analysis in the framework of autosegmental-metrical (AM) theory (e.g. [6]) is therefore an L* pitch accent.

Figure 1: Spectrogram and f_0 contour for male speaker BK.



The location of the f_0 peak is more variable: Keane [4] found inter-speaker differences and also differences between individual lexical items over whether the peak fell in the second or third syllable of trisyllabic words. Assuming that this reflects a phonological high tone, AM theory seems to offer two possible points of reference – the preceding low tone or the end of the word. In the first case, the high tone would form the trailing tone of a bitonal L*H pitch accent; in the second, it would be a boundary tone.

The current study was designed to investigate both the distribution of fall-rise-fall contours and the alignment of low and high turning-points in words of different lengths. The relation between the phonetic alignment of f_0 turning-points and the phonological association of pitch features is still not fully understood, despite intensive research in recent years (e.g. [7]). However, it seems reasonable to expect that the distance between the

low and high turning-points will increase as word length increases, if H is a boundary tone. At the same time, the distance between the peak and the end of the word should remain relatively constant. The consequences for alignment, should H be connected with the low tone, are less clear. According to Pierrehumbert [8], the unstarred tone in an American English L*H pitch accent trails its starred counterpart by *'a given time interval'* but subsequent work (e.g. [1, 2]) suggests that this claim is oversimplified. It appears likely, however, that the peak will be progressively further from the end of the word as word-length increases.

2. EXPERIMENT

2.1. Design

A set of target words was selected, such that both length (from monosyllabic through to tetrasyllabic) and syllable structure were systematically manipulated, as shown in table 1. As far as possible the words were fully voiced, although the absence of voiced word-initial obstruents from the native Dravidian vocabulary resulted in the inclusion of some voiceless word-initial stops. Concrete nouns were used so that each target word could be embedded in two carrier sentences, in phrase-medial and phrase-final position:

paɖatɨl oɾʉ ___ iɾʉndaðʉ
picture.loc a ___ be.past.3sn
'There was a ___ in the picture.'

iŋge: oɾʉ ___
here a ___
'Here is a ___.'

The lack of plurisyllabic monomorphemic concrete nouns in Tamil meant that the longer target words were inevitably compounds. Morpheme boundaries are marked in table 1 by hyphens or orthographic spaces.

Table 1: Target words.

Syllable quantities	Transcription	Translation
—	mi:n	'fish'
—	na:j	'dog'
∪ ∪	ʋaɻɨ	'path'
∪ ∪	naɾɨ	'fox'
— ∪	nuŋgʉ	'palmyra fruit kernel'
— ∪	ʋi:ɖʉ	'house'
∪ —	malaɪ	'mountain'
∪ —	ʋajal	'field'
— —	ma:ŋ-ga:j	'mango'
— —	ma:laɪ	'garland, necklace'
∪ ∪ ∪	miɻaɣʉ	'pepper'
— ∪ ∪	ka:j-ɣaɾɨ	'vegetable'
— — ∪	muɻ:aŋgɨ	'radish'
— — —	ʋeŋga:jəm	'onion'
— ∪ —	ma:-maɾəm	'mango tree'
∪ — ∪	kuɾaŋgʉ	'monkey'
∪ — —	ʋaɻaɪjəm	'bracelet, arm-ring'
∪ ∪ —	ʲeɾʉmaɪ	'buffalo'
— ∪ ∪ —	nel:i-maɾəm	'gooseberry tree'
— — ∪ —	koj:a:-maɾəm	'guava tree'
— — — ∪	ʋa:nəm-ba:ɖɨ	'skylark'
∪ — ∪ —	talaɪ-jaɳaɪ	'pillow'
∪ ∪ ∪ —	tiɾʉ-ʋiɻa:	'festival'
— ∪ ∪ ∪	me:l uðaɖʉ	'upper lip'
— ∪ — ∪	ma:ɖɨ ʋi:ɖʉ	'storeyed house'
— — — —	ten:əm-biɻ:aɪ	'coconut sapling'

Syllable quantity: — heavy ∪ light

2.2. Speakers

Recent studies of other languages (e.g. [3]) have led to an increasing awareness of cross-dialectal differences in intonation and, in the absence of specific data, it seems reasonable to assume that Tamil is affected by similar variation. This was therefore controlled as far as possible by choosing speakers that formed a relatively homogeneous group. Eighteen teenagers were recorded (9 male and 9 female), all of whom had spent their entire lives in Madurai, a city in central southern Tamil Nadu, and were thus unlikely to show regional variation. Age-related differences were also controlled by selecting subjects between 15 and 17-years-old. They attended schools that drew pupils from relatively affluent backgrounds, which should have reduced any variation correlated with social stratification.

The recordings took place in quiet rooms within the school complexes, using lapel microphones (Audio-technica AT803b) in conjunction with a portable CD recorder, and were digitised at a rate of 22 kHz (16 bit resolution). F_0 contours were

extracted from each of the 936 sentence tokens using 5ms windows on ESPS *xwaves*™ software.

2.3. Measurements

Measurements were made on simultaneous displays of waveform, wide-band spectrogram and f_0 track. Markers were placed at either end of the target word, using standard segmentation criteria. In a few cases the relevant boundary could not be located with confidence, so these were treated as missing data points.

In addition, all f_0 turning-points in the target word were marked by hand, and subsequently checked automatically against the f_0 values of the sample points on either side to ensure that they corresponded to local maxima and minima. Peaks were marked at absolute maxima except where a preceding voiceless stop induced a perturbation. Low turning-points proved much more difficult to identify consistently. In the end, the policy adopted was to work backwards from a peak and mark the first sample preceded by a higher f_0 value.

3. RESULTS

3.1. Distribution of fall-rise-fall contours

F_0 contours on the target words were of three types: a gradual decline, a single fall-rise-fall and a double fall-rise-fall. The relative distribution of these patterns (marked 'decline', 'single' and 'double' respectively in table 2) was influenced both by prosodic context and word structure. There was also considerable variation between speakers. For instance, male speaker PR had only one double contour anywhere, whereas there were 15 in the speech of female speaker SA.

As table 2 shows, the examples of gradual decline were all found in phrase-final target words, where it accounted for approximately half of the mono-, di- and trisyllabic tokens. The occurrence of this pattern in phrase-final position is therefore not restricted to verbs, but at the same time it is not obligatory for nouns, as it seems to be for verbs: the other half of the mono-, di- and trisyllabic tokens were marked by single fall-rise-falls.

In phrase-medial position mono-, di- and trisyllabic words were uniformly marked by single fall-rise-falls, with just three exceptional double contours. Two-thirds of phrase-medial tetrasyllabic words had double contours, the remaining third single ones. The double pattern was thus largely restricted to the longer target words. Neither length in terms of numbers of syllables nor morphological complexity was a necessary criterion for a double contour as there were exceptional cases of both shorter words and monomorphemes with this pattern. Conversely, there were many instances of tetrasyllabic words and dimorphemes with single fall-rise-falls, so the double pattern was not obligatory in these words. The extent to which length and morphological complexity co-occur in the data set is a problem for establishing which is more influential: tetrasyllabic monomorphemic nouns would be the obvious test case but such examples are lacking in Tamil.

3.2. Alignment

The first low turning-point in a phrase-medial target word occurred somewhere within the initial syllable in the vast majority of cases (465 out of 467). This confirms earlier work and supports an L* analysis. The alignment of the high turning-point showed considerable variability, but always within the bounds of the target word. In monosyllables it occurred towards the end of the single syllable; in disyllables it was found predominantly in the second syllable (132 out of 144). In trisyllabic words the position of the peak varied between medial and final syllables, as in [4], with approximately two-thirds in final syllables (91 out of 144) and a third in medial syllables.

As described, the original intention was to use words of increasing length to test whether the alignment of the high tone is tied to the preceding low or a following prosodic boundary. This was hampered by the tendency for longer words to bear double fall-rise-falls. Moreover, in cases where a single fall-rise-fall contour was found on a

Table 2: Distribution of f_0 patterns on target words

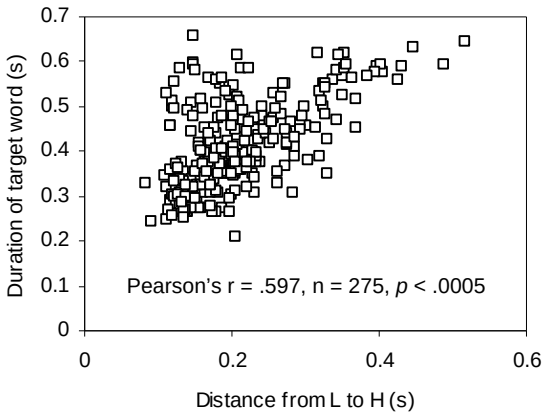
Target word	Phrase-medial position		Phrase-final position		
	Single	Double	Decline	Single	Double
Monosyllabic	100%	—	47%	53%	—
Disyllabic	100%	—	50%	50%	—
Trisyllabic	98%	2%	57%	42%	1%
Tetrasyllabic	34%	66%	10%	73%	17%
Monomorphemic	99%	1%	54%	46%	—
Dimorphemic	51%	49%	20%	66%	14%

tetrasyllabic (compound) word, the peak typically occurred in the second syllable, followed by a gradual decline on the second half of the word. There seems to be some restriction, therefore, on the distance over which a single fall-rise-fall contour can be realized. Prima facie this might appear to be evidence for the H tone being connected in some fashion with the L. However, the boundary within the compound provides an alternative point of association. H would then be a boundary tone marking the end of the prosodic constituent corresponding to the first element of the compound, at most a prosodic word.

Due to the uncertainty surrounding their significance, data from all the tetrasyllabic words and the two trisyllabic tokens with double contours were excluded from the statistical analysis. The prediction was that, if H is a boundary tone, there should be a clear correlation between the duration of the target word and the distance between the low and high f_0 turning-points. This was borne out to some extent: there was a significant correlation between the L to H distance and the duration of the target word as a whole (Pearson's r = .597, n = 275, $p < .0005$), but individual data points were quite widely dispersed, as the scatterplot in figure 2 illustrates.

Furthermore, the distance between H and the end of the target word, which is expected to be relatively constant for a boundary tone, varied considerably (mean 97 ms; std dev 64 ms). There is also a significant, although fairly weak, correlation between this distance and the duration of the target word (Pearson's r = .511, n = 276, $p < .0005$), the effect that was expected if the high tone were instead connected to the preceding low tone.

Figure 2: Scatterplot showing the distance between low and high f_0 turning-points against the duration of the target word.



4. DISCUSSION

The results of the investigation into alignment were not entirely conclusive. There is a significant correlation between the L to H distance and the duration of the whole word, matching the prediction for H as a boundary tone. However, this is not accompanied by the expected stability of H relative to the end of the word. There is also a correlation, albeit weaker, between this distance and the duration of the target word, which seems to point in the other direction, namely an association between L and H that might be evidence of a bitonal pitch accent. A third possibility, which is compatible with these findings but not easily accommodated by AM theory, is that the position of the high f_0 turning-point is fixed as a proportion of the target word. Further research is needed to distinguish definitively between these possibilities, although the balance of evidence favours the boundary tone analysis.

Any future investigation can benefit from an improved understanding of the distribution of fall-rise-fall contours in Tamil. In particular, it was established that a declining f_0 pattern phrase-finally is not specific to verbs, since it was found on some nominal tokens, but nor is it required by the prosodic environment, since others were marked by a fall-rise-fall. A second key finding was the occurrence of double fall-rise-falls on some longer words. Should H be a boundary tone, it must therefore mark the edge not of a phrase, but a lower level prosodic constituent, maximally the prosodic word.

5. REFERENCES

[1] Arvaniti, A., Ladd, D., Mennen, I. 2000. What is a starred tone? Evidence from Greek. In: Broe, M., Pierrehumbert, J. (eds), *Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press. 119-131.

[2] Dilley, L., Ladd, D., Schepman, A. 2005. Alignment of L and H in bitonal pitch accents: testing two hypotheses. *JPhon* 33, 115-119.

[3] Gilles, P., Peters, J. (eds) 2004. *Regional variation in intonation*. Tübingen: Niemeyer.

[4] Keane, E. 2006a. Prominence in Tamil. *Journal of the International Phonetic Association* 36, 1-20.

[5] Keane, E. 2006b. Phonetics vs. phonology in Tamil *wh*-questions. *Proc. Speech Prosody 2006* Dresden.

[6] Ladd, D. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

[7] Ladd, D. 2003. Phonological conditioning of F0 target alignment. *Proc. 15th ICPhS* Barcelona, 249-252.

[8] Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. diss., MIT.