Investigating the intonational phonology of Tamil*

Background

Tamil is a Dravidian language, spoken predominantly in the Indian state of Tamil Nadu. Last available Census of India figures (1991) indicate 53 million speakers within India; there are also large Tamil-speaking communities elsewhere, including Sri Lanka, Malaysia and Singapore.

Tamil has a long history as a written language, with a native tradition of grammatical analysis dating back at least as far as the early centuries AD (Tolkāppiyam). It was also one of the first South Asian languages to come to the attention of western scholars.

Intonation

Examples of typical f0 traces for declarative sentences:

Figure 1: Amplitude waveform and f0 contour

![Figure 1](image1.png)

Maadavi marundai vaanginaa.
Madavi medicine buy.past.3sf
‘Madavi bought medicine.’

Murugan epgaaoḍu Madurai-ilirundu varuvaam.
Murugan we.assoc Madurai.abl come.fut.3sm
‘Murugan will come with us from Madurai.’

Figure 2: Amplitude waveform and f0 contour

![Figure 2](image2.png)

Distribution of (fall-)rise-falls

Typically each word except the last in the phrase bears a (fall-)rise-fall but …

- some words simply have a gradually declining f0 trace, particularly function words, e.g. oru ‘a’ in figure 4
- morphologically complex words may have a double fall-rise-fall, e.g. nelli-maram ‘Indian gooseberry (type of tree)’ in figure 4, also Madurai-ilirundu ‘from Madurai’ in figure 2
- phrase-final nouns may bear a fall-rise-fall, e.g. valaiyam ‘bracelet’ in figure 5

*The support of a British Academy Overseas Conference Grant is gratefully acknowledged.
Phonological characterization of (fall-)rise-falls

A first pass: LH pitch accent assigned post-lexically (there are no minimal pairs distinguished by pitch, and the same lexical item can occur with and without a fall-rise-fall).

Position of lexical accent in Tamil?
- no minimal pairs distinguished by stress
- native speakers have no strong intuitions about which syllable, if any, is stressed (Balasubramanian 1980)

… but word-initial syllables in Tamil are special in some respects:
- evidence of weak phonetic prominence
  - qualitative vowel reduction – tokens of /a/, /i/ & /u/ have significantly different spectral characteristics from their counterparts in non-initial syllables (Keane 2003)
  - duration – some indications that vowels in initial syllables are longer than their counterparts in non-initial syllables (Keane 2003 but cf. Keane 2006b), but no evidence that the vowels of initial syllables are significantly louder (Keane 2006b, in line with Balasubramanian 1972)
- distributional asymmetries – greater range of vowel contrasts found in initial syllables than elsewhere in the word
The low $f_0$ turning-point occurs consistently within the word-initial syllable (Keane 2006b, 2007), so it seems reasonable to designate the low tone with a star, i.e. $L^*$.

Possible points of reference for the $H$ tone:

- preceding $L$ tone, i.e. $H$ would be the trailing tone of a bitonal $L^*H$ pitch accent
- the end of the word/morpheme, i.e. $H$ would be associated with the boundary of a low-level phonological constituent (maximally the prosodic word)

The predicted consequences for its phonetic alignment are not entirely clear: the suggestion that a trailing tone may be separated from the starred tone by ‘a given time interval’ (Pierrehumbert 1980) has been challenged (e.g. Arvaniti, Ladd & Mennen 2000; Dilley, Ladd & Schepman 2005).

Hypothesis under test (Keane 2007): if $H$ is a boundary tone, the distance between the low and high turning-points will increase as word length increases, cf. if $H$ is a trailing tone, the high turning-point will stay further forward in the word.

Results: not entirely conclusive, although more support for $H$ being a boundary than a trailing tone.

**Phrase-final intonation**

Typically the phrase-final constituent of a declarative has flattish or gradually declining $f_0$ (see figures 1, 2 & 4), consistent with an $L^\%$ boundary tone.

Clear final rises, reflecting an $H^\%$ phrase-boundary tone, are found in contexts including:

- continuation rises in the middle of an utterance, as indicated by the arrow in figure 6

Sentence offset $f_0$ is significantly higher in $wh$-questions than in the corresponding declaratives (Keane 2006a).

**Some outstanding issues**

- scaling – appears to play an important role in signalling relative semantic weight, e.g. $f_0$ peaks are raised in question words (Keane 2006a)
- focus – relative part played by intonation vs. focus particles
Methodology

Intonational database collected in the field in Tamil Nadu, modelled in part on the IViE database of British English dialects (http://www.phon.ox.ac.uk/IViE/).

Other languages have significant regional variation in intonation (Gilles & Peters 2004) and Tamil is probably the same, so speakers were selected to be maximally homogeneous.

Speakers:
- 24 (half male, half female)
- 15 to 17-years-old
- lifelong residents of Madurai (city in southern central Tamil Nadu)
- all spoke Tamil in the home
- recorded in same-sex pairs within schools

Database design:
- set of 137 sentences designed to investigate three issues:
  - alignment (Keane 2007)
  - interrogativity – wh- (Keane 2006a), echo, tag and yes-no questions (partially analysed)
  - focus (not yet analysed)
- short story (Teru Vilakkku ‘The street light’ by Pudumaippittan)
- questions and answers – translations of materials used in a crosslinguistic study of deaccenting (Cruttenden 2006)
- fable (Kaakkaavum Nariyum ‘The crow and the fox’)
- retelling of the fable
- Tamil version of the map task (Brown et al. 1983)
- free conversation on a given topic (marriage customs for the girls, cricketing heroes for the boys)

A tricky issue…

Diglossia – there are significant differences in syntax, lexis, morphology and even segmental phonology between formal and colloquial Tamil (e.g. Britto 1986). Contexts of usage include:
- formal Tamil – news broadcasts, church liturgy, public meetings, etc. and writing
- colloquial Tamil – everyday oral communication. There are no standards for writing it down.

The challenge: eliciting any kind of controlled speech in colloquial Tamil.

Previous (largely unsuccessful) tactics:
- present speakers with a text in standard spoken Tamil (Schiffman 1998) and tell them to use it only as a guide (Keane 2006c)
- ask speakers to respond to a pre-recorded question on the basis of a picture, thus avoiding a written representation altogether (Asher & Keane 2005)
More successful tactics:

• using younger speakers

• recording in a familiar environment, with the presence of a class-mate

• using a graded sequence of tasks, with those designed to elicit colloquial speech at the end
References


