Tamil

Elinor Keane
Oxford University Phonetics Laboratory
elinor.keane@phonetics.oxford.ac.uk

Tamil is a Dravidian language spoken by 53 million speakers in India, according to census figures from 1991, predominantly in the state of Tamil Nadu. There are also sizeable communities of Tamil speakers elsewhere, including Sri Lanka, Malaysia and Singapore, in all of which it has the status of a national language. Tamil is diglossic, the formal or 'literary' variety still largely conforming to standards set in the thirteenth century by the Tamil grammarian Pavanandi. It is used in almost all written media, and also for certain high-register functions. In all other situations colloquial Tamil is used and is characterized by considerable regional and social variation.

The recording on which the transcription is based was made by a female native speaker in her early fifties. She comes from near Madras and her speech is fairly representative of the variety sometimes known as 'Standard' Spoken Tamil (Schiffman 1998), which is derived from the speech of non-Brahmins from central Tamil Nadu and has gained wide currency in recent years. Spoken varieties of Tamil are rarely written down, and there are no standards for spelling, official or unofficial. The orthographic representation of the passage reflects some colloquial features, but is relatively conservative. The citation forms, by contrast, are more representative of the formal register, because a colloquial style is felt to be inappropriate for the careful pronunciation of individual words.

Consonants

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar
Plosive	p (b)		t̪ (d̪)			ʈ (ɖ)		k (g)
Affricate					tʃ (dʒ)			
Nasal	m			n		ɳ	(ɲ)	
Fricative				s				
Tap				ɾ				
Approximant		ʋ				ɻ	j	
Lateral approximant				l		ɭ		

Journal of the International Phonetic Association (2004) 34/1
DOI:10.1017/S0025100304001549

Obstruents

The Tamil orthography uses a single symbol at each place of articulation for all the obstruents, reflecting the fact that voicing is not contrastive in the native Dravidian vocabulary. Rather, there is complementary distribution of different phonetic alternants, sometimes referred to as Caldwell's Law (Caldwell 1856: 102). In word-initial position voiceless plosives are found: these may be accompanied by slight aspiration, although this seems to be variable. Voiced obstruents occur word-internally after nasal segments and involve complete occlusion of the oral tract. Intervocalically the exact realization depends on the place of articulation: for dentals there is variation between a voiced stop and fricative, and for bilabials there may be further weakening to an approximant. For retroflex sounds either a voiced stop or a flap is found between vowels. Several possibilities have been reported for the velar sounds, including a voiceless palatal fricative, voiced and voiceless velar fricatives, and the voiced glottal fricative [ɦ].

Intervocalically there is a contrast at each place of articulation between the various possibilities just described, which are represented by a single orthographic symbol, and a voiceless stop, which corresponds to an orthographic geminate. In phonetic terms the distinction between them involves several dimensions, including voicing and degree of occlusion. Duration may not be the primary factor, and there is conflicting evidence over whether word-initial and word-internal voiceless stops are consistently different in length (Balasubramanian & Asher 1984, Keane 2001). However, considerations of pattern congruity with nasals and laterals, as well as the orthography, support a phonological analysis of the word-internal voiceless stops as geminates.

Caldwell's Law holds only of a subset of the modern lexicon, and not the many loanwords that have entered Tamil from languages with distinctive voicing, primarily Indo-Aryan, Perso-Arabic and English. Consequently voiced plosives do occur in word-initial position, and these are included in the consonant chart, with their non-native origin signalled by parentheses.

In most spoken varieties of Tamil there is a five-way place-of-articulation distinction amongst the obstruents: bilabial, dental, postalveolar, retroflex and velar. One exception to this is the Kanniyakumari dialect, which distinguishes between dental and alveolar obstruents, realizing the orthographic geminate ற்ற as a voiceless alveolar plosive. The corresponding single symbol represents a rhotic sound, but in most dialects the geminate has merged with the voiceless plosive member of the dental series. Retroflex consonants do not appear in word-initial position in the native non-onomatopoeic lexicon. They are articulated with a high degree of retroflexion, certainly greater than that found in Hindi (Ladefoged & Bhaskararao 1983).

At the postalveolar place of articulation, a voiceless affricate tends to be found in word-initial position and as the intervocalic realization of a geminate, although a voiceless palatal stop or alveolo-palatal fricative are also possible in either position. As with the other series of obstruents, the voiced equivalent follows a nasal and can also occur word-initially in loanwords. The most common realization of a single intervocalic segment is the alveolar fricative [s], and this can also occur word-initially. In many instances the source is a loanword, but it is also found as a variant pronunciation (alongside [tʃ]) of a native ச. The relative distribution of initial [tʃ] and [s] is a matter of considerable variation: for some speakers it seems to be determined by the nature of the following vowel, but sociolinguistic factors are also involved. An initial [s] that is recognized as a borrowing may be represented orthographically as ஸ, while an initial [dʒ] is written as ஜ. These supplementary letters come from the Grantha writing system, which was used in south India for Sanskrit and in Sanskritized Tamil. Other fricatives that may be represented by Grantha letters are [ʂ] and [h]; [f] and [z] may also occur in loanwords but are relatively marginal, and thus none is included in the consonant chart.

Nasals

Tamil nasals occur contrastively in intervocalic position at three different places of articulation: bilabial, alveolar and retroflex. Word-initially there is a contrast between bilabial

https://doi.org/10.1017/S0025100304001549 Published online by Cambridge University Press

and alveolar nasals. The palatal nasal may also be found in this position but it is very rare in Standard Spoken Tamil, and hence its phonemic status is marginal. The vast majority of tokens of [ɲ] precede a homorganic affricate word-medially. A velar nasal [ŋ] is also found, but its distribution is completely predictable, occurring only before a single voiced velar obstruent. Corresponding geminates exist for the bilabial, alveolar and retroflex nasals (marginally for the palatal place of articulation) and, like other sonorant geminates in Tamil, are distinguished from their single counterparts by significantly greater durations (Balasubramanian 1982b). The alveolar nasal has two orthographic representations, ந and ன. These occur in near complementary distribution: ந is found almost exclusively word-initially and internally before a dental obstruent (in which case assimilation of place of articulation occurs), and ன elsewhere. However, this written distinction does not reflect any phonetic contrast in speech (Balasubramanian 1982a).

Liquids

There are two lateral approximants in Tamil, alveolar [l] and retroflex [ɭ], both of which can be geminated. Rhotic liquids, in contrast, are the only category of sounds in Tamil that do not undergo gemination. The orthography distinguishes between two, ர and ழ: each is subject to certain distributional restrictions, but in intervocalic position they may mark a lexical contrast. Opinion is divided over whether this corresponds to a phonetic distinction in colloquial Tamil: Balasubramanian (1982a) argues that ழ always corresponds phonetically to a tap, whilst [ɾ] and [r] are both possible realizations of ர, their distribution being determined in part by environment. Narayanan, Byrd and Kaun (1999) report a slight difference in place, rather than manner, of articulation, with ர pre-alveolar and ழ post-alveolar. In most colloquial varieties, however, it seems that a merger is in progress, if not complete.

A further liquid, orthographically ள, exists in many varieties of Tamil although production of the distinctive sound is sometimes restricted to formal speech. One possible substitute is [ɭ], to which it is acoustically very similar. Claims about its articulatory properties have been many and various, but two extensive recent studies, employing a range of experimental methods, have concluded that it is best described as a central retroflex approximant [ɻ] (McDonough & Johnson 1997 and Narayanan et al. 1999). The anterior tongue body is said to be curved up towards the central palatal region, with lateral contact between the sides of the mid-tongue and the palate.

In the following examples and the passage, the transcriptions are broadly phonetic, reflecting the distribution of the different phonetic alternants for the obstruents and also vowel reduction. The italicized words are transliterated from the Tamil orthography and can be regarded as largely phonemic.

p	paṯʉ	*pattu*	‘ten’	pː	apaː	*appaa*	‘father’
[b]	ṯambi	*tampi*	‘younger brother’	[ʋ]	koːʋəm	*koopam*	‘anger’
(b)	baɳɪl	*patil*	‘answer’				
ṯ	ṯapʉ	*tappu*	‘mistake’	ṯː	ṯaːṯaː	*taattaa*	‘grandfather’
[ɖ]	paɳɖʉ	*pantu*	‘ball’	[ð]	nalːəðʉ	*nallatu*	‘it is good’
(ɖ)	ɖinəm	*tinam*	‘day’				
ṭ				ṭː	paṭʉ	*paṭṭu*	‘silk’
[ɖ]	ʋanɖi	*vanṭi*	‘cart’	[ɖ]	paɖi	*paṭi*	‘step’
k	kaːl	*kaal*	‘leg’	kː	makəɭ	*makkaɭ*	‘people’
[g]	aŋgeː	*aŋkee*	‘there’	[ɣ]	maɣəɭ	*makaɭ*	‘daughter’
(g)	ganəm	*kanam*	‘heaviness’				
tʃ	tʃɪnːə	*cinna*	‘small’	tʃː	peːtʃːu	*peeccu*	‘speech’

https://doi.org/10.1017/S0025100304001549 Published online by Cambridge University Press

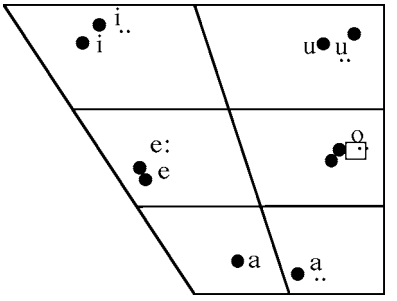
[dʒ]	ɪɲdʒi	*iɲji*	‘ginger’	[s]	peːsʉ	*peecu*	‘speak’
(dʒ)	dʒurəm	*juram*	‘fever’				
s	saːpɪɖʉ	*saappiɽu*	‘eat’				
m	aːmaːm	*aamaam*	‘yes’	mː	amːaː	*ammaa*	‘mother’
n	aːnaːl	*aanaal*	‘but’	nː	kanːəm	*kannam*	‘cheek’
ɳ	maɳi	*maɳi*	‘hour’	ɳː	aɳːən	*aɳɳan*	‘elder brother’
ɲ	aɲdʒi	*aɲci*	‘five’				
l	puli	*puli*	‘tiger’	lː	nalːə	*nalla*	‘good’
ɭ	puɭi	*puɭi*	‘tamarind’	ɭː	puɭːi	*puɭɭi*	‘dot’
ɾ	kaɾi	*kari*	‘charcoal’				
ɻ	vaɻi	*vaɻi*	‘way’				
ʋ	ʋali	*vali*	‘ache’	ʋː	oʊːoɳɖɾʉ	*ovvoɳɽru*	‘one by one’
j	jaːnaɪ	*yaanai*	‘elephant’				

Vowels

Phonologically Tamil distinguishes five different vowel qualities, and for each there is a contrast in length. The chart illustrates their relative positions for articulations in initial syllables: in non-initial syllables, /i/, /a/ and /u/ undergo reduction in both duration and quality. The close back vowel is fronted, and its degree of rounding decreases. It may undergo a partial merger with /i/ in non-initial syllables to [ʉ] or even [ɨ]. There are two diphthongs, /aɪ/ and /aʊ/, the second of which is restricted to only a few lexical items. The phonetic realization of /aɪ/ is again influenced by syllable position: in non-initial syllables its duration decreases, and its formant structure is also affected to the extent that it may be realized as [ɛ] or [a], depending on the dialect.

Vowel quality may also be altered by harmony effects: for instance, a short close vowel in an open syllable is lowered when the next syllable contains /a/ or /aɪ/. Rounding of front vowels preceding a syllable containing a rounded vowel is also possible in certain lexical items. Finally, some coarticulatory centralization of vowels occurs immediately preceding a retroflex consonant.

On-glides may be found before certain vowels in word-initial position: [ʲ] before /e(ː)/ and /i(ː)/ and [ʷ] before the two pairs of back vowels. Their occurrence is somewhat variable, but more likely when utterance-initial or following a pause. In connected speech, a word-final sequence of vowel and nasal is usually replaced by a nasalized vowel. In the case of the short vowels, this may be accompanied by a change in quality. In particular, there is a contrast between [ẽ], the realization of /an/, and a more retracted vowel, variously [õ], [ã] or [ɤ̃], the realization of /am/. One consequence of this nasalization is to increase the number of words that end phonetically in a vowel. In words which end in a liquid, the same effect may be achieved by insertion of a final epenthetic vowel (phonetically [ɨ]] or [ʉ]), thereby eliminating any consonant-final words from colloquial speech.



https://doi.org/10.1017/S0025100304001549 Published online by Cambridge University Press

a	pal	*pal*	‘teeth’	aː	paːl	*paal*	‘milk’
e	ʲeri	*eri*	‘burn’	eː	ʲeːɾʉ	*eeru*	‘climb’
i	nilәm	*nilam*	‘earth’	iː	niːlәm	*niilam*	‘blue’
o	ʷoʈʉ	*oʈʈu*	‘stick’	oː	ʷoːʈʉ	*ooʈʈu*	‘drive’
u	ʷuɾʉ	*uru*	‘shape’	uː	ʷuːɾ	*uur*	‘village’
aɪ	paɪ	*pai*	‘bag’	aʊ	ʋaʊʋaːl	*vauvaal*	‘bat’

Prosody

There are no lexical distinctions based on tone or stress in Tamil, and the location, and even existence, of stress are disputed. However, the pattern of vowel reduction in non-initial syllables indicates that word-initial syllables have some phonetic prominence. The main pitch movements in intonational phrases seem to be associated with initial and final syllables.

Transcription of recorded passage

ʷoɾʉ naːlʉ kaːʈ̃ũ suːɾijənʉ̃ ɾombə saɳdə poːʈikiʈʉ jaːɾʉ balәsaːlinːʉ paɳɖajʌ̃ poːʈaːŋgə ‖ kaːʈʉ solːitʃi ʈaːɳɖaːn balәsaːlinːʉ | suːɾijẽ solːitʃi ʈaːɳɖaːn balәsaːlinːʉ ‖ apə aɳɖə ʋaɻijaː oɾʉ ʋaɻipoːkәn kambәɻije poːʈikiʈʉ poːnãː ‖ kaːʈ̃ũ suːɾijənʉ̃ jaːɾʉ ʋaɻipoːkәne kambәɻije eɖʉkə ʋaɪkiɾaːŋgәloː aʊәŋgәɖaːn balәsaːlinːʉ paɳɖajʌ̃ poːʈaːŋgə ‖ moɖәlːe kaːʈʉ ʋeːgәmaː aɖitʃiðʉ | ɾombə balәmaː aɖikaːɖikə ʋaɻipoːkәn kambәɻije iɻʉʈʉ iɻʉʈʉ poːʈikiʈãː ‖ kaɖesile kaːʈʉ kambәɻijeːɖʉkə mudɻijaːmə kaɪʋiʈʉɾʉtʃi | aɖʉʈʉ suːɾijẽ ukiɾәmaː kaːɲdʒiðʉ | ʋejil aɖiɣәm aːgaːgə ʋaɻipoːkәn ʋepәm ʈaːŋgə mudɻijaːmə kambәɻije eɖʉʈʉ poːʈãː ‖ ʲiðe paːʈʉʈʉ kaːʈʉ suːɾijәɳɖaːn balәsaːli aðʉ paɳɖaɪʈɪle dʒaɪtʃʉɾʉtʃʉnːʉ oʈʉkiɖʉtʃi ‖

Orthographic version

ஒரு நாளு காத்தும் சூரியனும் ரொம்ப சண்டெ போட்டுகிட்டு யாரு பலசாலின்னு பந்தயம் போட்டாங்க. காத்து சொல்லுச்சு தான்தான் பலசாலின்னு. சூரியன் சொல்லுச்சு தான்தான் பலசாலின்னு. அப்ப அந்த வழியா ஒரு வழிப்போக்கன் கம்பளியெ போத்திகிட்டு போனான். காத்தும் சூரியனும் யாரு வழிப்போக்கனெ கம்பளியெ எடுக்க வைக்கிராங்களோ அவங்கதான் பலசாலின்னு பந்தயம் போட்டாங்க. மொதல்லெ காத்து வேகமா அடிச்சுது. ரொம்ப பலமா அடிக்க அடிக்க வழிப்போக்கன் கம்பளியெ இழுத்து இழுத்து போத்திகிட்டான். கடெசிலெ காத்து கம்பளியெ எடுக்க முடியாம கய்விட்டுருச்சு. அடுத்து சூரியன் உக்கிரமா காஞ்சுது. வெயில் அதிகம் ஆக ஆக வழிப்போக்கன் வெப்பம் தாங்க முடியாம கம்பளியெ எடுத்து போட்டான். இதெ பாத்துட்டு காத்து சூரியன்தான் பலசாலி, அது பந்தயத்திலெ ஜெய்ச்சுருச்சுன்னு ஒத்துகிடுச்சு.

https://doi.org/10.1017/S0025100304001549 Published online by Cambridge University Press

Acknowledgements

I would like to thank R. E. Asher and E. Annamalai for comments on an earlier draft of this paper and help with the passage, and also N. Kesavaram, the Tamil speaker.

References

BALASUBRAMANIAN, T. (1982a). The two r's and the two n's in Tamil. *Journal of Phonetics* **10**, 89–97.

BALASUBRAMANIAN, T. (1982b). Intervocalic double nasal and lateral consonant articulations in Tamil. *Journal of Phonetics* **10**, 99–104.

BALASUBRAMANIAN, T. & ASHER, R. E. (1984). Intervocalic plosives in Tamil. In Higgs, J.-A. W. & Thelwall, R. (eds.), *Topics in Linguistic Phonetics in Honour of E. T. Uldall* (Occasional Papers in Linguistics & Language Learning **9**), 49–63. Coleraine: The New University of Ulster.

CALDWELL, R. (1856). *A Comparative Grammar of the Dravidian or South-Indian Family of Languages*. London: Harrison & Sons.

KEANE, E. (2001). Echo words in Tamil. D.Phil. dissertation, University of Oxford.

LADEFOGED, P. & BHASKARARAO, P. (1983). Non-quantal aspects of consonant production: a study of retroflex consonants. *Journal of Phonetics* **11**, 291–302.

MCDONOUGH, J. & JOHNSON, K. (1997). Tamil liquids: an investigation into the basis of the contrast among five liquids in a dialect of Tamil. *Journal of the International Phonetic Association* **27**, 1–26.

NARAYANAN, S., BYRD, D. & KAUN, A. (1999). Geometry, kinematics, and acoustics of Tamil liquid consonants. *Journal of the Acoustical Society of America* **106.4**, 1993–2007.

SCHIFFMAN, H. (1998). Standardization or restandardization: the case for "Standard" Spoken Tamil. *Language in Society* **27**, 359–385.