# Language contact and sound change: Reasons for mutual unintelligibility between formal and colloquial registers of Tamil

ANUSHRI KARTIK-NARAYAN, *University of Colorado Boulder*

ABSTRACT

Tamil has since its origination been diglossic, separating the formal high register from the colloquial low register. These two registers are currently mutually unintelligible (Shanmugam Pillai 1965). This analysis explores the reasons why they became unintelligible, which are proposed to be two-fold: historic language contact between Tamil and Sanskrit; and sound changes demonstrated using the Comparative Method. It has been suggested that the decline in mutual intelligibility is due to the removal of Sanskrit loanwords from the formal high register during the Tamil Purist Movement of the 20th century (Kailasapathy 1979). The earliest evidence of Tamil and Sanskrit reciprocal borrowing dates to the first Tamil literary works (Krishnamurti 2003). Where and when this language contact occurred is unclear, but it may have occurred during overlapping occupation of the Indus River Valley region by Sanskrit and Proto-Dravidian (Steever 2009). During the 20th century, the formal register replaced these loanwords with Tamil equivalents wherever possible (Kailasapathy 1979). Currently, low register Tamil is composed of 50% loanwords whereas high register Tamil is composed of only 20% loanwords (Krishnamurti 2003). It has been attested, however, that some diglossia was present before contact between Tamil and Sanskrit. Early diglossia can thus instead be explained by sound changes, which also account for current differences between the registers not attributed to loanwords. Sound changes identified in this analysis include: syncope, apocope, paragoge, stop to fricative lenition, and others. This analysis finds that language contact and sound changes contributed to the decline in intelligibility between formal and colloquial Tamil, however the nature of the language contact is still under investigation.

## 1    Introduction

Tamil is a Dravidian language spoken in the southern third of the Indian peninsula as well as parts of Sri Lanka, Malaysia, and Singapore (Steever 2009). It is the most direct descendent of reconstructed Proto-Dravidian, which has been dated to circa 8,000 BCE. Ancient Tamil arose from Proto-Dravidian in 300 BCE, which then evolved into Medieval Tamil in 700 CE, and Modern Tamil in 1600 CE (Steever 2009).

Within Modern Tamil, multiple mutually intelligible dialects are divided by regional variety, caste, and even religious sect (*Vaishnavite*: those who worship Vishnu; *Shaivite*: those who worship Shiva) (Steever 2009). However, a divide between formal register Modern Tamil (hereafter referred to as "formal") and colloquial register Modern Tamil (hereafter referred to as "colloquial") varieties has created an unintelligibility that is only overcome by formal education of the formal variety and native acquisition of the colloquial variety. Native speakers of colloquial Tamil who do are not educated in formal Tamil cannot understand it, and people who are only educated in formal Tamil cannot understand colloquial Tamil (Shanmugam Pillai 1965). This phenomenon of formal and colloquial registers with different methods of acquisition, known as diglossia, is common to diverse languages globally, but is known to be mutually intelligible in

such cases (Ferguson 1959). The question this paper explores is thus, how did formal Tamil become unintelligible to colloquial Tamil?

## 2        Data analysis

In order to determine how colloquial Tamil became different from formal Tamil, the following data were analyzed following the Comparative Method of Sound Change. This data was sourced from ilearnTamil (2020), transcribed from Tamil script into the International Phonetic Alphabet (IPA) using Rajan (2014), and is presented here in tables with columns delineated English Gloss – Formal Tamil (Romanized) – Formal Tamil (IPA) – Colloquial Tamil (Romanized) – Colloquial Tamil (IPA). It is important to note that it is assumed that colloquial Tamil derives from formal Tamil, rather than both deriving from Proto-Dravidian simultaneously, and thus the proposed sound changes are in the direction of formal-to-colloquial. Additionally, the following sound changes are tendencies rather than rules because the changes do not always apply uniformly to a given context.

### 2.1        Syncope

Syncope is deletion of a phoneme or entire syllable word-internally. Syncope in Tamil can affect syllables such as [ɾi], [ɾʉ], and [gɪ]. This process is depicted by the data in Table 1.

### 2.2        Apocope

Apocope is deletion of a phoneme or entire syllable word-finally. The colloquial register of Tamil prefers words to end in vowels, with exceptions made for nasal consonants. Thus, non-nasal word-final consonants are deleted, such as [|] and [k]. This process is depicted by the data in Table 2.

### 2.3        Paragoge

Paragoge is the insertion of a phoneme word-finally. The colloquial register of Tamil inserts a vowel such as [ə] or [ʉ] word-finally due to the aforementioned preference for words to end in vowels. This process is depicted by the data in Table 3.

### 2.4        High front unrounded vowel becomes high not-front unrounded vowel

The vowel [ɪ] becomes either [ʉ] or [ʊ] as a backing process while maintaining height and rounding. This process is depicted by the data in Table 4.

### 2.5        Alveolar tap becomes velar nasal

The alveolar tap [ɾ] becomes the velar nasal [ŋ] preceding a velar stop as a partial assimilation process via the place-of-articulation feature. This process is depicted by the data in Table 5.

### 2.6        Vowel reduction / monophthongization

Word-final diphthongs either monophthongize to the first vowel in the segment or reduce to [ə]. This further contributes to the low register's preference for word-final monophthong vowels. This process is depicted by the data in Table 6.

## 2.7    Voiced coronal stop becomes voiceless fricative

Voiced coronal stops such as [d̪] and [ɖ] become voiceless coronal fricatives such as [s] as a lenition process where manner of articulation is weakened, and voicing is lost. Place-of-articulation is not always retained in the change. This process is depicted by the data in Table 7.

## 2.8    Exceptions to sound change processes

Not all differences between the formal and colloquial registers of Tamil can be explained with the aforementioned sound change tendencies, as depicted by the data in Table 8. The differences between these terms cannot be explained by sound changes and thus must be the result of borrowing through language contact. The most likely contact language candidate is Sanskrit, which will be discussed in the next section.

## 3    Contact with Sanskrit

Sanskrit is an Indo-European language that no longer has native speakers and whose current relevance is limited to historic texts and Hindu functions. It is derived from the Indo-Aryan branch of Proto-Indo-European. Vedic Sanskrit arose circa 1500 BCE, then evolved into Classical Sanskrit circa 700 BCE, and fell out of spoken use by 1350 CE (Steever 2009). Sanskrit's regional distribution, as attested by the controversial but still presently acknowledged Aryan Invasion Hypothesis, began in the Indus River Valley and entered the Indian peninsula circa 2000 BCE (Steever 2009). Language contact may have occurred within the Indian peninsula during the origination of Tamil from Proto-Dravidian, however alternate theories suggest contact in the Indus River Valley during trade. The true temporal and spatial nature of contact between Tamil and Sanskrit is still under investigation.

Sanskrit literary scholars are known to have interacted and collaborated with Tamil *sangams* (literary scholars) from the beginning of Tamil's literary tradition, leading to reciprocal borrowing between the two languages. (Krishnamurti 2003). The oldest known Tamil text, the *tolkappiyam* grammar of Tamil, contains borrowed terms from Sanskrit (Krishnamurti 2003). Following a trace of the number of Sanskrit words in Tamil literature, the highest increase in proportion occurred during Medieval Tamil from 300-600 CE (Krishnamurti 2003). Borrowing of Sanskrit words was common in all of the Dravidian languages of the area, and it continued well into modern forms of these languages (Krishnamurti 2003). However, at the turn of the 20th century a surge in Tamil nationalism led to a drive for Tamil revivalism (Kailasapathy 1979). Scholars actively attempted to recreate a "pure" Tamil free of borrowed terms from Sanskrit. This movement was somewhat successful and led to a reduction in Sanskrit loanwords from 50% of formal Tamil to 20% in present day, leaving only particularly abstract concepts and proper names of religious figures (Krishnamurti 2003). The high proportion of loanwords was retained in colloquial Tamil, creating a strong separation between the two registers.

## 4    Conclusion

Tamil's formal and colloquial registers both utilized loanwords from Sanskrit to an equal extent until the purist movement of the 20th century. However, diglossia has been attested as early as the literary tradition began during the period of Ancient Tamil. It is therefore possible that until the

20th century, the two registers were mutually intelligible, and the removal of Sanskrit loanwords from formal Tamil reduced mutual intelligibility. This would better fit the understood state of global diglossia (Ferguson 1959), but may force a reckoning of whether formal and colloquial Tamil may still be considered registers or should instead be considered a reconstructed language and a mixed language.

## Acknowledgements

## References

Ferguson, Charles A. 1959. Diglossia. Word 15(2). 325-340. DOI: 10.1080/00437956.1959.11659702

ilearnTamil. 2020. Spoken tamil lessons. https://ilearntamil.com/spoken-tamil-lessons-colloquial-tamil-greetings-and-introduction-lesson1/ (11 December, 2022).

Kailasapathy, K. 1979. The tamil purist movement: A re-evaluation. Social Scientist 7(10). 23-51. DOI: https://doi.org/10.2307/3516775

Krishnamurti, B. 2003. The Dravidian Languages. 1st edn. Cambridge University Press.

Shanmugam Pillai, M. 1965. Merger of literary and colloquial tamil. Anthropological Linguistics 7(4.II). 98-103. DOI: https://www.jstor.org/stable/30022544

Steever, Sanford B. 2009. Tamil and dravidian languages. In Comrie, Bernard (ed.), The World's Major Languages. 2nd edn. 635-652. New York, NY & Canada: Routledge.

Rajan, V. 2014. Tamil transcription tool. https://anunaadam.appspot.com/ (12 December, 2022)