1. Introduction

This document proposes an additional documentation to the Unicode Standards Version 4.0, section 9.6, *Ligatures*, item 1. In particular, it mentions the occurrence of conjunct கை as well as non-conjunct கை in modern Tamil. In general, Tamil avoids the use of conjunct letters and this phenomenon is unique among Indian languages. The Grantha conjunct கை is written as a non-conjunct கை in a wide variety of word categories in modern Tamil. It is important to note that Indian languages other than Tamil always employ கை in conjunct form. On the contrary, major classes of borrowed words in Tamil from Urdu or Arabic or English use the non-conjunct கை as the norm.

Even Sanskrit words (written in Tamil script) with கை (0b95 0bcd) followed by இ (0bb7) are written as non-conjunct கை for which examples are shown in Section 3.0 of this document. While other Indian languages use the conjunct கை exclusively, the pattern of using non-conjunct கை is more common in Tamil. In fact, in contemporary Tamil printed books, the frequency of non-conjunct கை is larger than that of the conjunct கை. Hence, this document recommends:

a. Creating the conjunct கை using a zero width joiner (ZWJ, U+200D) only when needed by the user;
   Changing the current use of கை, இ, இ sequence to generating only the non-conjunct form கை, and,

b. Changing the Named Sequence definition for TAMIL LETTER KSSA accordingly to:

   `<U+0B95, U+0BCD, U+200D, U+0BB7> TAMIL LETTER KSSA`

   (from the current Unicode consensus definition `<U+0B95, U+0BCD, U+0BB7> TAMIL LETTER KSSA`)

   Note: Use of கை, இ, இ (with ZWNJ) to produce the non-conjunct form கை though will be redundant, will not be affected by this proposal.

2. Proposed change to text in Section 9.6

Following text should be added to Section 9.6, subsection *Ligatures*, item 1:

In Tamil, unlike other Indian languages possessing conjuncts as their norm, both non-conjunct and conjunct forms are used whenever கை is followed by Grantha இ. These non-conjunct கை and conjunct இ forms can be treated together for collation of Tamil Grantha letters. Indeed, the non-ligating கை is found widely in modern Tamil publications. A zero width joiner (ZWJ, U+200D) should be placed between the characters கை and இ when the conjunct form இ is needed.

Conjunct Consonants:

\[\text{த + ஓ + [ZWJ] + இ} \rightarrow \text{தை (kṣā)}\]

In the case of conjunct letter, vowel rendering occurs around it. For example,

\[\text{த + ஓ + [ZWJ] + இ} + \text{ஓ + ப} \rightarrow \text{தைஹப (kṣā)}\]
3. Usage examples:

Example 1:
Dr. R. Kalaikkovan, Thalicceri Kalvettu, South Indian Saiva Siddhanta Publishing Society, Chennai, 2002. Note that the word, rāksai ‘benediction’ is written as conjunct ரகசை on page 115 and its equivalent form as non-conjunct ரகசை on page 258.

Page 115

1. குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு
2. பார்வை பார்வை பார்வை பார்வை
3. குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு
4. குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு
5. குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு
6. குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு
7. குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு
8. குறிப்பிட்டு குறிப்பிட்டு குறிப்பிட்டு


ligate rāksai
Non ligate
rakṣa
Example 2:
Dr. S. S. Janaki, Siva temple and temple rituals, Kuppuswami Sastri Research Institute, University of Madras, 1988.
In page 188, the Sanskrit loan word, ukṣar ‘Vedic deity’ is printed in non-conjunct form.
Example 3

In page 56, Sanskrit word akṣa 'eye' is printed in non-conjunct form.
Example 4
Normally, nowadays loan words from Western and Islamic world into Tamil are written as non-conjunct.
Following pages are from Jeyakanthan, Sila nerangalil sila manitharkal, Madurai, 1977.

Page 89, sekšan = ‘section’

Page 101, rikšaa = ‘rickshaw’