

Comparative Constructions in Tamil and English- A Comparative Study with the existing Machine Translation systems

ABSTRACT

Comparative construction is a linguistic technique for demonstrating two or more items to show their similarities and differences. Typically, a comparative construction consists of a predicate and two noun phrases, one of which serves as the "criterion" of the comparison and the other as the object of comparison (the comparee NP). The sentences that are comparable to sentences like "Raja is taller than Roja", where the noun phrase after the item "than" is the standard NP, are prototype examples of comparative constructions in the languages of the globe. As suggested by Dixon [3], a prototypical comparative construction has three elements: participants of comparison (comparee and standard of comparison), the property (parameter of comparison), and the index of comparison. In this paper, we are not concerned with the typological study of comparative constructions. We have focused on mapping comparative constructions in Tamil with that of English and formulated rules for Machine Translation. The translation between Comparative construction sentences in Tamil and English are tested with the existing Machine Translation system and their results are discussed.

Keywords: *Comparative Construction, Tamil language, comparee, Standard of comparison, Degree Marker, Comparison of Inequality, Comparison of Equality, Machine Translation*

1. INTRODUCTION

A prototypical comparative construction involves a quality or property whose extent is compared, the entity being compared, and the standard of comparison. Comparison is a mental act by which two or more items are examined in order to assess similarities or differences between them. The comparison can be made with regard to a certain gradable, one-dimensional property, and the items are then assigned a position on a predicative scale. This mental act of comparison finds its linguistic encoding in comparison constructions, especially comparative constructions for the expression of comparison of inequality or equative constructions for the expression of comparison of equality. In this research, we have tested and compared the grammatical structures found in the Tamil language with those present in English [6-8]. This extensive analysis has enabled us to establish a set of rules that can be applied to facilitate Machine Translation between the two languages. To validate the effectiveness of our findings, we conducted thorough testing of the translation process for comparative construction sentences in both Tamil and English using an established Machine Translation system. The outcomes of these tests have been meticulously analyzed and are presented for discussion and evaluation.

2. LITERATURE REVIEW

The linguistic literature has especially been concerned with comparison of inequality and comparative constructions as found in the following English sentences [5].

(1) Mary is tall-er than Peter -Comparee - Parameter - Parameter/Degree Marker - Standard Marker - Standard

(2) Mary is more intelligent than Peter- Comparee - Parameter/Degree Marker - Parameter - Standard Marker - Standard

- Comparee (COM) = what is being compared against some standard of comparison (*Mary*); alternative terms used in the literature: Item compared.
- Standard of comparison (SOC) = what the comparee is being compared against (*Peter*)
- Standard Marker = marker of the grammatical function of the standard (*than*), alternative terms used in the literature: marker, pivot, relator
- Parameter of comparison (POC) = property of comparison (*tall, intelligent*); alternative terms used in the literature: Quality or Quantity, comment, (comparative) predicate.
- Degree Marker or Parameter Marker: It marks the degree of presence or absence of a property in the comparee (*more* or *-er*); alternative terms used in the literature: Index, comparative concept.

Dixon [1] makes use of the terms comparee, index, parameter, mark and standard. The following example will exemplify the use of these terms.

John **is more** famous than Bill.

comparee index **parameter mark** standard

In traditional grammar of English, four degrees of comparison of the adjective are distinguished. They are positive degree, equative **degree, comparative** degree, and superlative degree [5].

Positive degree: basic form of the adjective; *Susan is tall* - **positive** construction

Equative degree: parameter is ascribed to the comparee and the standard to the same extent; *Susan is as tall as Peter* - equative construction

Comparative degree: Parameter applies **to the** comparee to a higher extent than to the standard; *Susan is tall-er than Peter* - comparative construction

Superlative degree: shows the highest degree of the parameter applied to the comparee; *Susan is the tallest of her family* - **superlative** construction

Tamil does not **makes** use of degree marker or parameter marker; it makes use of only parameter of comparison. In Tamil, there is no morphological distinction between positive degree, comparative degree and superlative degree. Rajendran [4] elaborately studied comparison of inequality and equality in Tamil.

3. COMPARISON OF INEQUALITY

In a prototypical comparative construction in Tamil and English, the comparee occupies the subject position, and the standard of comparison occupies the predicate position. In English, the standard of comparison occupies a position at the end of the comparative construction after the parameter of comparison whereas in Tamil, the standard of comparison occupies the predicative position before the parameter of comparison. The standard of comparison is

marked for accusative case in Tamil. The parameter of comparison does not make use of a comparative degree marker; it is pronominalized to agree with the subject NP. English makes use of *than* as parameter of comparison and Tamil makes use of *viTa* or *kaaTTilum* as parameter of comparison.

(3) raaNi raataiy-ai vita/kaaTTilum azahkaana-val

Rani Radha-ACC POC ADJ-PN

' Rani is more beautiful than Radha '

The mapping between Tamil and English can be given as follows:

Mapping rule 1:

$NP_{COM} + NP_{ai} + vita/kaaTTilum + ADJ-PN = NP_{COM} + BE + more + ADJ + than + NP_{soc}$

In the superlative comparative construction, the parameter of comparison is marked for superlative degree in English. In Tamil, the parameter of comparison is not marked for superlative degree. The standard of comparison has to be an inclusive NP of superlative nature meaning 'of all', 'among all' and so on. In Tamil too, the standard of comparison in superlative comparative construction must be inclusive nature: *avarkaL elloorilum* 'among all'.

(4) raaNi avarkaL elloor-ai-yum viTa ahakaana-val

Rani they all-ACC-EMP than ADJ-PN

Rani is most beautiful among all

Mapping rule 2:

$NP + NP_{ai} + viTa/kaaTTilum + ADJ-PN = NP_{COM} + BE + most + ADJ + among all.$

In English, certain adjectives inflect for comparative degree is marked with -er instead of *more*.

(5) raaNi raataiy-ai viTa/kaaTTilum uyaramaana-val

Rani Radha-ACC POC ADJ-PN

Rani is taller than Radha

Mapping rule 3:

$NP_{COM} + NP_{ai} + vita/kaaTTilum + ADJ-PN = NP_{COM} + BE + ADJ-er + than + NP_{soc}$

Some adjectives in English, inflect for comparative marker -est instead of **most** .

(6) raaNi avarkaL elloor-ai-yum viTa uyaramaana-val

Rani they all-ACC-EMP than tall-PN

Rani is tallest among all

Mapping rule 4:

NP+ NP-ai + viTa/kaaTTilum + ADJ-PN = NP_{COM} + BE + ADJ-est + among all.

In the place of ADJ-PN Tamil can make use of noun denoting quality +adverbial marker when followed by the be-verb *iru*. English makes use of be-verb and adjective combination only. The following sentence will exemplify this statement.

(7) raaNi raataiyai viTa/kaaTTilum azhak-aaka iru-kkiR-aaL

Rani Radha-ACC than beauty-ADVP be-PRE-3FS

Rani is more beautiful than Radha

Mapping rule 5

NP_{COM} + NP-ai + vita/kaaTTilum + N-ADVP iru-TEN-PNG= NP_{COM} + BE + more + ADJ + than + NP_{SOC}

Similar to adjectives, adverbs too can make comparison of inequality. Consider the following example.

(8) raaNi raataiy-ai viTa/kaaTTilum veekamaaka ooT-in-aaL

Rani Radha-ACC than fast run-PAS-PNG

Rani ran faster than Radha

In Tamil, the standard of comparison is marked for accusative marker; the standard marker *viTa/kaaTTilum* comes next and the parameter of comparison which is an adverb comes after standard marker and before the verb. In English, the verb comes before the parameter of comparison which is an adverb which is inflected for the comparative marker -er; the standard of comparison comes at the end and the standard marker *than* comes before the standard of comparison.

Mapping rule 6:

NP_{COM} + NP-ai + viTa/kaaTTilum + ADV +V-TEN-PNG = NP_{COM} + V-TEN + ADV-er + than + NP_{SOC}

4. COMPARISON OF EQUALITY

If two or more items are found to be similar quantitatively or qualitatively they can be subjected to the comparison of equality. Consider the following sentence.

(9) raaNi raataiy-aip poola azhakaana-vaL

Rani Radha-ACC like ADJ-PN

'Rani is as beautiful as Radha'

In the above sentence, Rani is the comparee, i.e. item compared; Radha is the standard of comparison; *poola* is the standard marker; *andazhakaana* 'beautiful' is the parameter of comparison. In English, the standard of comparison comes at the end; the parameter of comparison comes in-between the standard marker as-----as.

Mapping rule 7:

NP + NP-ai + poola + ADJ-PN = NP + BE + as-ADJ-as + NP_{SOC}

Similar to adjectives, adverbs also undergo comparison of similarity. Consider the following sentence.

(10) raaNi raataiy-aip poola veekamaaka ndaTa-kkiR-aaL

Rani Radha-ACC like walk-PRE-3FS

'Rani walks as fast as Radha'

In Tamil, the subject function as the comparee; the standard of comparison marked for accusative case follows it; the standard marker *poola* follows next; the parameter of comparison which is an adverb follows it and the verb which inflect for tense and person-number-gender (PNG) occupies the final position of the construction. In English, the subject NP function as the comparee. The verb which is inflected for tense comes next. The parameter of comparison comes in between the standard marker 'as---as'.

Mapping rule 8:

NP_{COM} + NP-ai + poola + ADV + V-TEN-PNG = NP_{COM} + V-TEN + as-ADV-as + NP_{SOC}

Comparison can be made without the explicit expression of adverb. In that context, the comparative construction become ambiguous. Consider the following example:

(11) raaNi raataiy-aip poola ootu-kiR-aaL

Rani Radha-ACC like run-PRE-3FS

'Rani runs like Radha'

The Tamil sentence is ambiguous as it can be interpreted in a number of ways: 'Rani, runs (instead of walking) like Radha', 'Rani runs in the same speed like Radha', 'Rani runs in the same style or manner like Radha' and so on. Consider the following example,

(13) raaNi-kku raataiy-aip poola ceelai iru-kkiR-atu

Rani-DAT Radha-ACC like sari be-PRE-3NS

'Rani has sari like Radha'

The Tamil sentence is ambiguous inviting different interpretations: Rani has similar sari like Radha, Rani has sari of same colour like Radha, Rani has sari of same texture like Radha, and so on.

Comparison can be made without explicitly expressing the parameter of comparison. Consider the following example:

(14) raaNi raataiy-aip poola iru-kkiR-aaL

Rani Radha-ACC like be-PRE-3FS

Rani resembles Radha.

The not-expression of parameter of comparison makes this sentence ambiguous allowing different interpretation from the point of view of quality and quantity.

Mapping rule 9:

NP_{COM} + NP-ai + poola + iru-TEN-PNG = NP_{COM} + resemble-TEN + NP_{SOC}

iru can be replaced by *toonRu* 'appear' in the above construction of equality.

(15) raaNi raatay-aip poola toonRu-kiR-aaL

Rani Radha-ACC like appear-PRE-3FS

'Rani appears like Radha'

The addition of emphatic -ee can make the resemblance more closer.

(16) raaNi raataiy-aip poolav-ee iru-kkiR-aaL

Rani Radha-ACC like-EMP be-PRE-3FS

'Rani resembles Radha very much'

Instead of *poola* 'like', *maatiri* 'like', *aLavukku* 'as much', *attanai* 'that many' can be made use of as standard marker.

(17) raaNikku raataiy-ai maatiri pasi.

Rani-DAT Radha-ACC like hunger

'Rani is hungry like Radha'

(18) raaNikku raataiy-ai aLavukku pasi.

Rani-DAT Radha-ACC that much hunger

'Rani is that much hungry like Radha'

(19) raaNikku raataiy-ai attanai pasi.

Rani-DAT Radha-ACC that much hunger

'Rani is as much hungry as Radha'

aLavU and *attanai* specifies quantity. Another way of expressing quantity for the sake of comparison is using *ettanai* 'how much' and *attanai* as exemplified in the following comparative construction.

(20) raaNi-kku ettanai ceelai iru-kkiR-at-oo attanai ceelai raatai-kk-um iru-kkiR-atu

Rani-DAT how-much be-PRE-3NS-Q that-much saree Radha-DAT-EM be-PRE-3NS

'Radha has as many as saris Rani'

ettanai---attanai, *evvaLavU---avvaLavU* can be equated with *as many as* and *as much as* respectively. Similarly *eppaTi---appaTi* and *evvaaRu--avvaaRu* can be equated with English *what manner--that manner* as exemplified in the following example.

(21) raaNi eppaTi ooT-in-aaL-oo appaTi raat-aiy-um ooT-in-aaL

Rani how run-PAS-3FS-Q that-manner Radha-ACC-EMP

Rani waked in the same manner like Radha'

Exact resemblance can be expressed by making use of the emphatic markers *-ee* and *taan* as exemplified by the following sentence.

(22) raaNi raatai-ee taan

Rani Radha-EMP EMP

'Rani is exactly like Radha'

camamaaka 'equally', *iNaiyaaka* 'equally' can be used to specify the exactness in the resemblance.

(23) raaNi raatai-kku samamaaka/iNaiyaaka paaT-in-aaL

Rani Radha-DAT equally sing-PAS-3FS

'Rani sang as equally as Radha'

One can notice similarity expressed in the following comparative phrases. In these constructions *poonRa* is used as standard marker.

(24) mati poonRa mukam

moon like face

'moon like face'

(25) taamarai poonRa mukam

lotus like face

'lotus like face'

Comparison of equality can be studied elaborately, but such an elaborative method is not adopted here.

5. COMPARATIVE STUDY WITH THE EXISTING MACHINE TRANSLATION

Comparative construction Sentences in Tamil are translated to English using the existing Translation Systems. The table given below shows the translation rules for the Comparative construction Sentences between Tamil and English and also the results of the existing Translation Systems.

Table 1. Comparative Study with the existing Machine Translation

Rule No	Tamil sentence & Rule	English Translation & Rule	Google	Bing - Microsoft	Systran
1.	ராணிராதாவைவிட அழகானவள் (NP _{COM} + NP-ai + vita/ kaaTTilum + ADJ-PN)	Rani is more beautiful than Radha (NP _{COM} + BE + more + ADJ + than + NP _{SOC})	Rani is more beautiful than Radha Yes	Rani is more beautiful than Radha Yes	Queen is better than Radha No
2.	ராணி அவர்கள் எல்லோரையும் விட அழகானவள் (NP+ NP-ai + viTa/kaaTTilum + ADJ-PN)	Rani is the most beautiful among all (NP _{COM} + BE + most +	Rani is more beautiful than all of them	The queen is more beautiful than all of them	The Queen is more beautiful than them all

		ADJ + among all)			
3.	ராணிராதாவைவிட உயரமானவள் (NP _{COM} + NP-ai + vita/kaaTTilum + ADJ-PN)	Rani is taller than Radha (NP _{COM} + BE + ADJ-er + than + NP _{SOC})	Rani is taller than Radha	Rani is taller than Radha	Queen is taller than Radha
4.	ராணி அவர்கள் எல்லோரையும் விட உயரமானவள் (NP+ NP-ai + viTa/kaaTTilum + ADJ-PN)	Rani is the tallest among all (NP _{COM} + BE + ADJ-est + among all)	Rani is taller than all of them	The Queen is taller than all of them	The Queen is taller than them all
5.	ராணிராதையைக்காட்டிலும் அழகாக இருக்கிறாள் (NP _{COM} + NP-ai + vita/ kaaTTilum + N-ADVP iru-TEN-PNG)	Rani is more beautiful than Radha (NP _{COM} + BE + more + ADJ + than+ NP _{SOC})	Rani is more beautiful than Radha	Rani is more beautiful than Radha	Queen is more beautiful than queen

6.	<p>ராணிராதாவைவிடவேகமாக டினாள்</p> <p>(NP_{COM} + NP-ai +viTa/ kaaTTilum + ADV +V-TEN-PNG)</p>	<p>Rani ran faster than Radha</p> <p>(NP_{COM} + V- TEN+ ADV-er +than + NP_{SOC})</p>	<p>Rani ran faste r than Radh a</p>	<p>Rani ran faster than Radha</p>	<p>Queen ran faster than Radha</p>
7.	<p>ராணிராதாவைப்போல் அழகான வள்</p> <p>(NP + NP-ai + pool + ADJ-PN)</p>	<p>Rani is as beautif ul as Radha</p> <p>(NP + BE + as-ADJ- as + NP_{SOC})</p>	<p>Beau tiful like Rani Radh a</p>	<p>Rani is as beauti ful as Radha</p>	<p>She is as beautif ul as queen Radha</p>
8.	<p>ராணிராதாவைப்போல் வேகமாக நடக்கிறாள்</p> <p>(NP_{COM} + NP-ai + poola + ADV + V- TEN-PNG)</p>	<p>Rani walks as fast as Radha</p> <p>(NP_{COM} + V-TEN + as- ADV-as +NP_{SOC})</p>	<p>Rani walks as as fast as Radh a</p>	<p>Rani walks as fast as Radha</p>	<p>Queen acts swiftly as Radha</p>
9.	<p>ராணிராதாவைப்போல் இருக்கி றாள்</p> <p>(NP_{COM} + NP-ai + poola + iru-TEN- PNG)</p>	<p>Rani looks/ resembl es Radha.</p> <p>(NP_{COM} + Look/ resembl</p>	<p>Look s like Rani Radh a</p>	<p>Rani looks like Radha</p>	<p>She is like Rani Raddh a</p>

		e-TEN + NP _{soc})			
10	ராணிராதையேதான்	'Rani is exactly like Radha'	Rani is Radha	It's Rani Radha	Rani Raddha
	Correct output percentage		60%	70%	40%

*The table is good, but need to be revised- especially we miss the mark on the translation. Then how output is calculated?.

6. CONCLUSION

The Comparison of equality and inequality needs elaborate study. Only certain important aspects of comparison are studied here from the point of view of machine translation. We try to map comparative constructions of equality and inequality in Tamil and English by posing mapping rules. The Comparative construction Sentences in Tamil are translated to English using the existing Translation Systems and the results indicate that the existing Translation system needs to concentrate on the specific pattern of comparative construction Sentences between Tamil and English. The Proposed Mapping rules will enhance the results of the existing Machine Translation systems.

REFERENCES

- [1] R.M.W. Dixon, Comparative constructions in English, Studia Anglica Posnaniensia, 41, 2005.
- [2] R.M.W.Dixon, Comparative constructions: A cross linguistic typology, Studies in Languages 32,4:787-817, 2008.
- [3] R.M.W.Dixon, Basic Linguistic Theory, Volume 3, Further Grammatical Topics. Oxford: Oxford University Press.7, 2012.
- [4] S.Rajendran, Comparative constructions in Tamil, Bulletin of the Deccan College Research Institute, volume xxxvi nos 1-4, 1976
- [5] Yvonne Treis, Comparative constructions: An Introduction. Linguistic Discovery volume 16, issue 1, 2018.

6. Murugaiyan A. Focus constructions in modern Tamil. *International Journal of Dravidian Linguistics*. 2009 Jun 1;38(2):43-69.
7. Thennarasu S. The Construction "I Love You" Comparative Accounts in Tamil and English. *Language in India*. 2011 Aug 1;11(8).
8. Eldho E, Kumar R. Choice of language in the construction of cultural identity by Tamil speakers in India. *International Journal of Language and Culture*. 2023 Dec 31;10(1):54-86.

*See the journal guidelines.

UNDER PEER REVIEW