ISOLATING WORD LEVEL RULES IN TAMIL LANGUAGE FOR EFFICIENT DEVELOPMENT OF LANGUAGE TOOLS

Suriyah M, Aarthy Anandan, Anitha Narasimhan and Madhan Karky
Karky Research Foundation, India

ABSTRACT

With the advent of social media, the amount of text available for processing across different natural languages has become enormous. In the past few decades, there has been tremendous increase in the number of language processing applications. The tools for natural language computing of various languages are very different because each language has its own set of grammatical rules. This paper focuses on identifying the basic inflectional principles of Tamil language at word level. Three levels of word inflection concepts are considered – Patterns, Rules and Exceptions. How grammatical principles for word inflections in Tamil can be grouped in these three levels and applied for obtaining different word forms is the focus of this paper. These can be made use of in a wide variety of natural language applications like morphological analysis, morphological generation, word level translation, spelling and grammar check, information extraction etc. The tools using these rules will account for faster operation and better implementation of Tamil grammatical rules referred from [தல்தேப்பியம் | tholgaappiyam] and [னன்று | nannool] in NLP applications.

KEYWORDS

Natural language processing, Rule based approach, word level rules, Tamil tool, language tools

1. INTRODUCTION

The number of social media platforms for various purposes like blogging, micro-blogging, photo sharing, social networking etc., has risen steadily. The number of people who are members of the same has also increased. For instance, the micro-blogging website Twitter has about 330 million users (on 1st January 2018) and about 500 million tweets (on 24th January 2017) are sent on an average per day [1]. This in turn, has paved way for collection and processing of various forms of data – text, image, video etc., generated with high velocity and veracity. Focusing on the processing of text, the posts of users in different natural languages have to be analysed; this avenue of NLP has many challenges. One of them is to analyse text and interpret the meaning conveyed. This is not possible without a tool which has knowledge of all the grammatical rules followed in the natural language considered. Different tools need to be used for processing different natural languages because of the differences that exist in terms of the syntax and grammatical rules. In this work, a three level grouping of grammatical principles for word inflections in Tamil language is proposed for better grammatical analysis.
Inflection in grammar refers to the modification of a root word to convey various characteristics like tense, gender, number, person, case etc. Inflection expresses one or more grammatical categories with a prefix, suffix or infix, or another internal modification such as a vowel change [2].

For example, consider the Tamil word [கிளிஞ்சல் | kiLinjalgaL | sea shell <plural>]. Here, [கிளிஞ்சல் | kiLinjal | sea shell] is the root word which is a noun in singular form. The suffix [ள் | kaL |] has been added to obtain the inflected version of the word in the plural form.

[கிளிஞ்சல் தளொல் | kiLinjalgaLaal | sea shell <plural><instrumental case>] is another inflected form of the same noun where the root word கிளிஞ்சல் is added with two suffixes – [ள் | kaL |] and [ஆல் | aal |] to denote plural and instrumental case respectively.

[கிளிஞ்சல் ளிடமிருந்து | kiLinjalgaLidamirundhu | sea shell <plural><locative case><ablative case>] is yet another form of the same root word [கிளிஞ்சல் | kiLinjal | sea shell] with three suffixes. The suffixes are [ள் | kaL |] denoting plural, [இடம் | idam |] denoting locative case and [இருந்து | irundhu |] denoting ablative case.

**EXAMPLE 1: SAMPLE NOUN INFLECTIONS**

1.1 கிளிஞ்சல் ள் ← கிளிஞ்சல் <root word> <noun> + கள் <plural>  
   kiLinjalgaL ← kiLinjal + kaL  
   sea shell<plural> ←  sea shell + plural suffix

1.2 கிளிஞ்சல் தளொல் ← கிளிஞ்சல் <root word><noun> + கள்<plural> + ஆல்<instrumental case>  
   kiLinjalgaLaal ← kiLinjal + kaL + aal  
   sea shell <plural><instrumental case> ←  sea shell + plural suffix + instrumental case suffix

1.3 கிளிஞ்சல் ளிடமிருந்து ← கிளிஞ்சல் <root word> <noun> + கள் <plural> + இடம் <locative case> + இருந்து <ablative case>  
   kiLinjalgaLidamirundhu ← kiLinjal + kaL + idam + irundhu  
   sea shell<plural><locative case> <ablative case> ← sea shell + plural suffix + locative case suffix +ablative case suffix

**Example 2: Sample verb inflections**

2.1 பாடுவாய் ← பாடு <verb> + வ் <future tense> + அய்<second person> <singular>  
   paaduvaay ← paadu + v + aay  
   sing <imperative> ← to sing + future tense + second person + singular

2.2 பாடுகொண்டுசூரு ← பாடு <verb> + இக்கொண்டுசூரு <past tense marker> + கள்<sandhi> + கொண்டுசூரு <auxiliary verb> + இக்கொண்டுசூரு <past tense marker> + அய்<third person femine singular suffix>  
   paadikkkoNdirundhaaL ← paadu + i + k + koNdiru + nth + aal  
   she was singing ← to sing + past continuous + third person + singular + feminine
In this work, three levels of word inflection rules are considered – Rules, Patterns and Exceptions. How grammatical rules for word inflections in Tamil can be grouped in these three levels and applied for obtaining different word forms is the focus of this paper. For the sake of simplicity, in this paper we call a morphological ending, a pattern. A Pattern can be a number or case ending associated with a noun or a person-number-gender marker associated with a verb. We have a set of Rules and Exceptions associated with each pattern. A Rule will define how the pattern can add to a root word or how it can separate from a morphological variation to obtain the root. An Exception to a rule will handle special cases where the rule cannot be applied directly. These three levels of grammatical forms discussed in this paper facilitate simple and efficient implementation of grammatical rules in basic NLP tools like morphological analyser, morphological generator, word level translator etc.

2. LITERATURE REVIEW

The three main types of morphological relations are Inflection, Derivation, and Compounding. Inflectional morphology modifies the properties of lexemes, while maintaining the basic meaning of the lexeme [3].

Singh and Sarma aim at reducing the number of affixes and word generation and re-adjustment rules describing nominal inflection by using inflectional class information to analyse various forms of Hindi nouns. They propose that the inflectional behaviour of nominal forms should be used entirely in the formation of nominal classes in Hindi [4].

To Ramscar, it appears that inflection is carried out through analogical reminding based on semantic and phonological similarity and that a rule-based route is not necessary to account for past tense inflection [5].

Agglutination is a process of morphology in which complex words are formed by joining morphemes together without changing the spelling or phonetics [6]. Prominent south Indian languages like Tamil, Telugu, Malayalam, Kannada and Tulu are agglutinative [7].

Reddiah suggests that in addition to having similar sounds, all Dravidian languages are related closely in grammatical aspects too. There are similarities in prefix, suffix, numbers, gender and syntax and that all the Dravidian Languages are formed from one main Dravidian language, which may be Tamil by most scholars’ opinions [8].

Anand kumar et al propose a machine learning based approach for morphological analysis using SVM. This is a supervised model which needs voluminous training data to get good results [9]. But rule based approaches do not have this overload of having enormous tagged training data. TamilComputingLab morphological analyser is a conventional rule based one [10]; Piripori is the Tamil morphological analyser built on the principles discussed in this paper [11]. It was observed from the working of the above two morphological analysers that when morphological rules are
grouped and streamlined following the approach mentioned in this work, the speed of processing
and the accuracy with respect to precision and recall are increased compared to the conventional
rule based approach. The grouping of inflectional rules in Tamil accounts for faster and more
accurate operation of language tools; hence becoming necessary for better tools and applications.

3. **Description Of Inflectional Rules In Tamil**

Tamil script has 12 vowels [யுிரீ தமிழ் தொட்டு குக்கை | uyirezhukkaL], 18 consonants
[மேயெயே தமிழ் தொட்டு குக்கை | meyyeyezhukkaL] which combine to form 216 combinational
letters [யுிரீமேயே தமிழ் தொட்டு குக்கை | uyirmeyyeyezhukkaL] summing up to 247 letters.
Tamil has more than 4 lakh words. These words are classified as [பெய் சொல் | peyachol | noun], [வினா சொல் | vinaichol | verb], [இடைச் சொல் | idaichol | preposition and post
position], [உரிச் சொல் | urichol | adjective], [இலக்கியச் சொல் | ilakkiyachol | literary
words]. Apart from the classes mentioned, Tamil has the ability to adapt words from other
languages; these words are called [வடசொல் | vadasol | words adapted from Sanskrit],
[திலசொல் | thisaichol | word adapted from other languages like Urdu, Portuguese,
Telugu etc] [12]. These words combine with various prefixes and suffixes to yield their different
inflectional forms. When these combinations happen, certain alphabet transformations may (or
may not) happen. This work focuses on the collection of different prefixes and suffixes added to
Tamil root words (which are nouns or verbs) and the different grammatical transformations which
may happen during inflection. [தொல்லத்தியம் | tholgaappiyam] and [நன்புல் |
nannool] are the sources of the inflectional rules discussed in this paper.

4. **Proposed Work**

4.1 **Patterns**

These are suffixes added to root words to get the different morphological variations of the words
considered. These patterns contain morphological information of number and case for nouns;
tense, number, case, person and gender for verbs.

Example 3 (Inflectional forms of a noun): The Tamil root word, [மலர் | malar | flower], has
been taken as a sample to demonstrate some of the inflectional forms of nouns in Table 1. [மலர்
| malar | flower] is a noun whose inflected forms are obtained by mere suffixing of appropriate
patterns. Some other nouns may require some transformation, addition or removal of certain
letters. This is discussed in the Rules section.
Table 1. Different inflectional forms of the noun [மலர் | malar | flower].

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Implication of the pattern</th>
<th>Inflectional split of the pattern</th>
<th>Inflected form of the noun</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[கள்</td>
<td>kaL ]</td>
<td>&lt;plural&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[மலர் + கள்</td>
<td>malar + kaL ]</td>
</tr>
<tr>
<td>[களின்</td>
<td>kaLin ]</td>
<td>&lt;plural&gt; &lt;ablative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களால்</td>
<td>kaLaal ]</td>
<td>&lt;plural&gt; &lt;instrumental case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களை</td>
<td>kaLai ]</td>
<td>&lt;plural&gt; &lt;accusative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களிலிருந்து</td>
<td>kaLilirundhu ]</td>
<td>&lt;plural&gt; &lt;ablative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களியாய்</td>
<td>kaLidam ]</td>
<td>&lt;plural&gt; &lt;locative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களியாய்</td>
<td>kaLidamirundhu ]</td>
<td>&lt;plural&gt; &lt;ablative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களலுக்</td>
<td>kaLukku ]</td>
<td>&lt;plural&gt; &lt;dative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களுடன்</td>
<td>kaLudan ]</td>
<td>&lt;plural&gt; &lt;associative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
<tr>
<td>[களில்</td>
<td>kaLil ]</td>
<td>&lt;plural&gt; &lt;locative case&gt;</td>
<td>[மலர்</td>
<td>malar</td>
</tr>
</tbody>
</table>

Example 4 (Inflectional forms of a verb): The Tamil root word, [பொர் | paar | see], has been taken as a sample to demonstrate different inflectional forms of verbs when affixed with different patterns denoting number and cases.

The noun and verb considered in Table 1 and Table 2 are such that their inflectional forms can be obtained by mere suffixing of the patterns given. But there are other words which would require...
 transformation of one or more alphabet to obtain the inflected form when combining with a pattern. Such examples are discussed in the “Rules” part.

Table 2. Different inflectional forms of the verb [பொ | paar | see]

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Implication of the pattern</th>
<th>Inflectional split of the pattern</th>
<th>Inflected form of the verb</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[பொ</td>
<td>paay</td>
<td>]</td>
<td>&lt;second person&gt; &lt;singular&gt; &lt;future tense&gt;</td>
<td>ப்பொ (சுற்றி) + ப்பொ (சுற்றி) + + ப்பொ (சுற்றி) சுற்றி ) + ப்பொ (சுற்றி) சுற்றி ) +</td>
</tr>
<tr>
<td>[பொ</td>
<td>ppaargal.]</td>
<td>&lt;third person&gt;</td>
<td>&lt;plural&gt; &lt;future tense&gt;</td>
<td>ப்பொ (சுற்றி) + ப்பொ (சுற்றி) + ப்பொ (சுற்றி) சுற்றி ) + ப்பொ (சுற்றி) சுற்றி ) +</td>
</tr>
<tr>
<td>[பொ</td>
<td>ppaal.]</td>
<td>&lt;third person&gt;</td>
<td>&lt;plural&gt; &lt;future tense&gt;</td>
<td>&lt;female&gt;</td>
</tr>
<tr>
<td>[பொ</td>
<td>ppeer ]</td>
<td>&lt;second person&gt;</td>
<td>&lt;plural&gt; &lt;future tense&gt;</td>
<td>ப்பொ (சுற்றி) + ப்பொ (சுற்றி) + ப்பொ (சுற்றி) சுற்றி ) + ப்பொ (சுற்றி) சுற்றி ) +</td>
</tr>
<tr>
<td>[பொ</td>
<td>ppaar ]</td>
<td>&lt;third person&gt;</td>
<td>&lt;plural&gt; &lt;future tense&gt;</td>
<td>ப்பொ (சுற்றி) + ப்பொ (சுற்றி) + ப்பொ (சுற்றி) சுற்றி ) + ப்பொ (சுற்றி) சுற்றி ) +</td>
</tr>
<tr>
<td>[பொ</td>
<td>paen ]</td>
<td>&lt;first person&gt;</td>
<td>&lt;singular&gt; &lt;future tense&gt;</td>
<td>ப்பொ (சுற்றி) + ப்பொ (சுற்றி) + ப்பொ (சுற்றி) சுற்றி ) + ப்பொ (சுற்றி) சுற்றி ) +</td>
</tr>
<tr>
<td>[பொ</td>
<td>ppaom ]</td>
<td>&lt;first person&gt;</td>
<td>&lt;plural&gt; &lt;future tense&gt;</td>
<td>ப்பொ (சுற்றி) + ப்பொ (சுற்றி) + ப்பொ (சுற்றி) சுற்றி ) + ப்பொ (சுற்றி) சுற்றி ) +</td>
</tr>
<tr>
<td>[பொ</td>
<td>ppaan ]</td>
<td>&lt;third person&gt;</td>
<td>&lt;singular&gt; &lt;future tense&gt;</td>
<td>&lt;masculine&gt;</td>
</tr>
<tr>
<td>[பொ</td>
<td>peergal.]</td>
<td>&lt;second person&gt;</td>
<td>&lt;plural&gt; &lt;future tense&gt;</td>
<td>ப்பொ (சுற்றி) + ப்பொ (சுற்றி) + ப்பொ (சுற்றி) சுற்றி ) + ப்பொ (சுற்றி) சுrif</td>
</tr>
</tbody>
</table>
4.2. RULES

Each pattern has a set of Rules associated with it which records the list of all possible transformations which may occur when that pattern combines with different words to form the inflected forms of the same.

When a root word combines with a specific pattern (suffix with some morphological information), the former may have a change in one or more letters, or may have one or more letters added to it to form the final inflected form. These transformations are governed by grammatical principles described in [13], [14] and [15]. These are classified as “Rules” under each pattern i.e., each pattern has an associated set of Rules governing the transformations which may occur when the pattern combines with a word to form the inflected form of the latter.

The relationship between patterns and rules is illustrated in the Figure 1.

Let

\[ n \text{ - number of patterns} \]
\[ \text{Rules}(x) \text{ - the set of rules for the pattern ‘x’} \]
\[ \text{Rule } x,y \text{ - Rule number ‘y’ of the pattern ‘x’} \]
\[ \text{np } x \text{ - number of rules for pattern ‘x’} \]

Each of the ‘n’ patterns have a set of rules associated with it. A pattern is a suffix which can be affixed with a word to form its morphological variants. Each of the rules under a pattern are numbered with a unique number also denoting the pattern it is associated with.

Fig 1. Relationship between patterns and rules
Example 5 (Rules associated with a noun pattern): The following table illustrates some of the rules which are associated with the pattern [ள் | kaL] which denotes plural.

Table 3. Rules associated with the pattern [ள் | kaL]

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Root noun</th>
<th>Rule</th>
<th>Inflected form</th>
<th>Inflectional split</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ள்</td>
<td>kaL]</td>
<td>[மரம்</td>
<td>maram</td>
<td>tree]</td>
</tr>
<tr>
<td></td>
<td>[வொழ்</td>
<td>vaazhthu</td>
<td>greeting]</td>
<td>[k &gt; k]</td>
</tr>
<tr>
<td></td>
<td>[முள்</td>
<td>muL</td>
<td>thorn]</td>
<td>[t &gt; t]</td>
</tr>
<tr>
<td></td>
<td>[சொல்</td>
<td>chol</td>
<td>word]</td>
<td>[R &gt; 1]</td>
</tr>
</tbody>
</table>

In Table 3, the noun [மரம் | maram | tree], has [m] as the last letter. When [மரம் | maram | tree] is pluralized with the suffixing of the pattern [ள் | kaL], the “m” of “மரம்” has to be transformed to [ng | ng]. [மரம் | maram | tree] has a [k] attached to it when pluralized with the suffixing of the pattern [ள் | kaL]. [முள் | muL | thorn] ends with the alphabet [L | l]. When suffixed with the pattern [ள் | kaL], the “L” of “முள்” changes to [t | l].

Example 6 (Rules associated with a verb pattern): Coming to Rules for verb inflections, some of the rules of the pattern [பட்டிரு ் | pattirukka] when combining with verbs are studied.

Table 4. Rules associated with the pattern [பட்டிரு ் | pattirukka]

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Root verb</th>
<th>Rule</th>
<th>Inflected form</th>
<th>Inflectional split</th>
</tr>
</thead>
<tbody>
<tr>
<td>[இது</td>
<td>do]</td>
<td>[y &gt; y]</td>
<td>[தசய்யப்பட்டிரு ்</td>
<td>cheyyappattirukka</td>
</tr>
<tr>
<td>[வழங்கு</td>
<td>vazhangu</td>
<td>give]</td>
<td>[k &gt; k]</td>
<td>[வழங்கப்பட்டிரு ்</td>
</tr>
<tr>
<td>[வுடு</td>
<td>moodu</td>
<td>close]</td>
<td>[t &gt; t]</td>
<td>[வுட்பட்டிரு ்</td>
</tr>
</tbody>
</table>
In Table 4, the root words end with the letters [கு | ku], [சு | chu], [து | tu], [து | thu], [பு | pu], [று | Ru] which are formed by the combination [வல்லின | ezhuthu] with the vowel [இ | i]. The vowel [இ | i] changes to [அ | a] in these transformations.

### 4.3. Exceptions

These are transformations which do not follow any regular structure with respect to the patterns discussed earlier.

Example 7 (Some sample exceptions for verbs): The following table illustrates some exceptions which occur with verbs.

<table>
<thead>
<tr>
<th>Root verb</th>
<th>Pattern</th>
<th>Implication of the pattern</th>
<th>Inflected form of the word</th>
<th>Intermediate inflectional split</th>
<th>Inflectional split</th>
</tr>
</thead>
<tbody>
<tr>
<td>வொ</td>
<td>vaa</td>
<td>come</td>
<td>[கண்கன</td>
<td>ndhaan</td>
<td>&lt;third person&gt;</td>
</tr>
<tr>
<td>வொ</td>
<td>vaa</td>
<td>come</td>
<td>[கெமன</td>
<td>kiRaan]</td>
<td>&lt;third person&gt;</td>
</tr>
<tr>
<td>நில்</td>
<td>nil</td>
<td>stand</td>
<td>[சுணை</td>
<td>aan]</td>
<td>&lt;third person&gt;</td>
</tr>
</tbody>
</table>
The verb [வொ | vaa | come] combines with the pattern [ந் | ndhaan] to form [வந் | vandhaan| he came] i.e., the alphabet [வொ | vaa] changes to [வ | va]. The same verb combines with the pattern [கி | kiRaan] to form [வருகி | varugiRaan| he comes]. In this inflection, the alphabet [வொ | vaa] changes to [வ | va] and the alphabet [ரு | ru] gets added as well.

5. RESULTS AND DISCUSSION

280 noun patterns and 880 verb patterns were collected. The number of rules for nouns, verbs and compound words are 50, 124 and 54 respectively. Each pattern is given a specific ID. All the rules applicable for a specific pattern are grouped with the ID allotted for the pattern. These inflection rules form a repository which can be used to form a variety of NLP tools like morphological analyser, morphological generator, word level translator, spelling and grammar checker etc.

The grouping and classification of inflectional rules in Tamil discussed in this paper have been applied to build Piripori [11], a Tamil morphological analyser. A list of 23292 Tamil words were taken from different sources and were input to Piripori and Tamil Computing Lab analyser. In Table 6, the analysis times for 137144 Tamil words for Tamil Computing Lab analyser [10] and Piripori [11] are tabulated. The time taken by Piripori to analyse the 23292 words given as input is lesser than that taken by Tamil Computing Lab analyzer. The graphical representation of the data is given in Figure 2.

<table>
<thead>
<tr>
<th>Analysis time in nanoseconds</th>
<th>Analysis time in minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TaCoLa analyser</strong></td>
<td>4.6553E+11</td>
</tr>
<tr>
<td><strong>Piripori</strong></td>
<td>44597119480</td>
</tr>
</tbody>
</table>
Fig 2. Comparison of analysis times of Tamil Computing Lab analyser and Piripori

To compare precision and recall of the two analysers, a standard POS tagged document [16] was taken from LDCIL website as test data. The file consisted 164 words out of which 121 were nouns and 43 were verbs. Precision and recall of analysis of noun and verbs in the abovementioned file were calculated separately. The results of the same are tabulated in Table 7.

Table 7. Data for precision and recall calculation

<table>
<thead>
<tr>
<th></th>
<th>Tacola</th>
<th>Piripori</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of nouns correctly classified</td>
<td>85</td>
<td>110</td>
</tr>
<tr>
<td>No. of verbs correctly classified</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>No. of nouns classified as verbs</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>No. of verbs classified as nouns</td>
<td>12</td>
<td>5</td>
</tr>
</tbody>
</table>

Precision and Recall are calculated using the following formulae.

Precisionnoun = (No. of nouns correctly analysed) / (No. of words analysed as nouns)

Precisionverb = (No. of verbs correctly analysed) / (No. of words analysed as verbs)

Recallnoun = (No. of nouns correctly analysed) / (No. of nouns)

Recallverb = (No. of verbs correctly analysed) / (No. of verbs)

The results of the calculation are tabulated in Table 8 and represented graphically in Fig 3.

Table 8. Precision and Recall of nouns and verbs

<table>
<thead>
<tr>
<th></th>
<th>TaCoLa</th>
<th>Piripori</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision of Nouns</td>
<td>0.87628866</td>
<td>0.9565217</td>
</tr>
<tr>
<td>Recall of Nouns</td>
<td>0.702479339</td>
<td>0.9090909</td>
</tr>
<tr>
<td>Precision of Verbs</td>
<td>0.689655172</td>
<td>1</td>
</tr>
<tr>
<td>Recall of Verbs</td>
<td>0.465116279</td>
<td>0.5813953</td>
</tr>
</tbody>
</table>
The precision and recall values of nouns and verbs of Piripori, which is based on the grouping of grammatical rules discussed in this paper are higher than those of Tamil Computing Lab analyser.

Thus with the application of the grouping of grammatical rules into Patterns, Rules and Exceptions as discussed in this paper has led to a better performing Tamil morphological analyser in terms of analysis time, precision and recall compared to another analyser following the traditional rule based approach.

![Comparison of Precision and Recall](image)

**Fig 3. Comparison of Precision and Recall**

6. **CONCLUSION**

Three levels of Tamil word inflection rules are considered in this work – Rules, Patterns and Exceptions. How grammatical rules for word inflections in Tamil can be grouped in these three levels and applied for obtaining different word forms is the focus of this paper.

Efficient application of syntax and grammatical rules of a natural language in a natural language application or tool is very important for minimal error. This work focuses on the collection of inflectional rules in Tamil grammar referring [13], [14] and [15] thus forming a repository which can be used in various NLP tools and applications. There has been great improvement in retrieval speed and accuracy with respect to precision and recall when these rules were used to build a morphological analyser compared to the traditional rule based approach.

**ACKNOWLEDGEMENTS**

The authors would like to thank Subalalitha C.N, Kanniammal.R, Manju Bashini P, Elanchezhiyan K, Rajapandian C, Swathi P.G and Sindhuja G for their guidance and contribution to this work.
REFERENCES


AUTHORS

Suriyah M is a post graduate in Computer Science and Engineering with 4 years’ experience in teaching graduate engineering students. Now she is a research writer with Karky Research Foundation.

Aarthy Anandan is a post graduate in Computer Applications. Now she is working as Software Developer with 4 years’ in Karky Research Foundation.

Anitha Narasimhan is a post graduate in Computer Applications. Now she is working as Software Developer with 4 years’ in Karky Research Foundation.

Madhan Karky is a Ph.D in Computer Science. He is the Founder and Research Head of Karky Research Foundation, a nonprofit educational research organization which focuses on language computing and language literacy.