ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 12(29), DOI: 10.17485/ijst/2019/v12i29/147082, August 2019*

Tamil Spell Checker App for iPhone

S. Thendral[1*], R. Subhashini[1] and V. Madhan Karky[2]

[1]School of Computing, Sathyabama Institute of Science and Technology, Chennai - 600119, Tamil Nadu, India; thendraldots@gmail.com, subhaagopi@gmail.com
[2]Karky Research Foundation, Chennai - 600024, Tamil Nadu, India; inbox@karky.in

Abstract

Objective: To obtain an application for iPhone mobile to identify and correct the spelling mistakes in the given Tamil text and post it in the social media. **Methods**: The application splits each Tamil word from the given Tamil text and further analysed by using Tamil dictionary. Each word is split morphologically and finally checked for its sandhi errors based on Tamil grammar rules. **Findings**: Currently, Tamil spell check application is not available for iPhone in app store. So, this will be first application for identifying misspellings in Tamil text. **Applications/Improvements:** This Tamil spell checker application can be improved by adding grammar rules for other languages too.

Keywords: iPhone, Tamil Spell Checker, NLP

1. Introduction

Mobile phones are taking important role on media with memes, comments and reviews. This application can correct the spelling mistakes in Tamil text and post their corrected output in twitter, Facebook and google plus. This application developed based on Tamil dictionary words, morphological generator, sandhi checker and grammar rules in Tamil books like Nannul and Tholkappiam. While using this iPhone application, user can validate their content, and this can correct words automatically[1,2].

2. How It Develops

The spell checker develops by the following steps:

- Dictionary analysis
- Morphological analysis
- Sandhi analysis

2.1 Analyse in Dictionary

The application using web API to check the words in agaraadhi dictionary which develops and maintains by Karky research foundation. It contains more than 2.5 lakh Tamil words. Thus, the identified root words from the agaraadhi are either noun or verb.

2.2 Morphological Analysis

The application splits the word morphologically as paghuthi, idainilai, vighuthi, sandhi and sariyai. The paghuthi is the root word which can't further be split. It classified as peyarpaghuthi (noun) and vinaipaghuthi (verb). The idainilai shows the tense past (ith, it, ir, in), present (kiru, kindru, aninru) or future (ip, iv). The vighuthi shows gender (un, aan, ul, aal, ir, kal) and count (thu, na, in, il, kal). The algorithm developed based on

*Author for correspondence

the nannool (Tamil grammar book) grammar rules. The application checks the following for noun and verb:

- Plural - The root word + kal or mar or margal.

For example: amma + kal = ammakkal, amma + mar = ammamar,amma + margal = ammamargal.

- Vetrumaiurubu- case suffixes like ai, aal, ku, in,athu, kan added to the words.

For example: amma + ai = ammavai, amma + aal = ammaval, amma + in = ammavin, amma + odu = ammavodu,amma + udan = ammavudan, amma + ilirunthu = ammavilirunthu, amma + udaiya = ammavudaiya, amma + idamirunthu = ammavidamirunthu, amma + il = ammavil, amma + idam = ammavidam.

- Ottu–hooker word like a,e,o,than,mattum added to the words

For example: amma + a = ammava, amma + e = ammave, amma + o = ammavo, amma + than = ammathan, amma + mattum = ammamattum, amma + mathiram = ammamathiram, amma + ennum = ammavennum, amma + akilum = ammavakilum, amma + ayinum = ammavayinum, amma + um = ammavum, amma + avathu = ammavavathu, amma + ada = ammavada, amma + adi = ammavadi, amma + amma = ammavamma, amma + appa = ammavappa, amma + ayya = ammavayya

- Panmaiottu – combination of plurals and hooker words.

word + kal / mar / margal + a/ e/ o/ than/ mattum/ mathiram/ ennum/ akilum/ ayinum/ um/ avathu/ ada/ adi/ amma/ appa/ ayya

For example: ammakkale, ammakala, ammamar galmattum,ammamargalayinum

- Panmaivetrumaiurubu– combination of plurals and case suffixes.

word + kal / mar / margal + ai/ aal/ ku/ in/ odu/ udan/ ilirunthu/ udaiya/ idamirunthu/ il/ idam/ athu/ kan

For example: ammakkalai, ammakkallal, ammakkalin, ammakkalodu.

- Vetrumaiurubuottu– combination of case suffixes and hooker words.

word + case suffixes + a/ e/ o/ than/ mattum/ mathiram/ ennum/ akilum/ ayinum/ um/ avathu/ ada/ adi/ amma/appa/ ayya

For example : ammavidamiruntha, ammavilirunthe, ammavaiyo, ammavirka.

- Vetrumaiurubusollurubu – combination of case suffixes and word suffixes.

word + case suffixes + ida/ idavum/ pol/ pondru/ kondu/ nokki/ patri/ kurithu/ sutri/ sutrilum/ vittu/ thavira/ munnittu/ vendi/ otti/ poruthu/ poruthavarai/ agha/ endru/ mun/ pin/ ul/ idaiye/ naduve/ mathiyil/ mel/ kizh/ ethiril/ pakkathil/ arugil/ pathil/ maraga/ neraga/ uriya/ ulla/ thakuntha/ vayilaka/ moolamaka/ vazhiyaka/ peril/ poruttu/ kooda/ udaiya/ vasam/ varai/ thorum.

For example: ammavaipol, ammavaipondru, ammavaikondu, ammavainokki, ammavaipatri, ammavaikurithu, ammavaisutri, ammavaivittu

- Vetrumaiurubusolurubuottu – combination of case suffixes, word suffixes and hooker words.

word + case suffixes + word suffixes + hooker words.

For example: ammavaividavuma, ammavaimunnitte, ammavinmoolamakavo, ammavaividamattum, ammavinporuttayinum, ammavukkethirilennum.

2.3 Sandhi Analysis

If checking sandhi for two words, while initial word ending in vowel and next word starting in hard consonant ka,sa, tha, pa then hard consonants doubles in the following places:

The first word is a single letter word, adverbial participle, hard consonant shortened u,medium consonant shortened u, property compound, clitic word, cardinal numerals eight and ten, place names, indirect simile, 4th and 6th indirect cases, 2nd,3rd,4th,5th,7th urubumpayanumudanthokkathogai and metaphor.

3. How to use Application

Type up to 1000 words at a time for check in input window. The iPhone application analyzes each word and

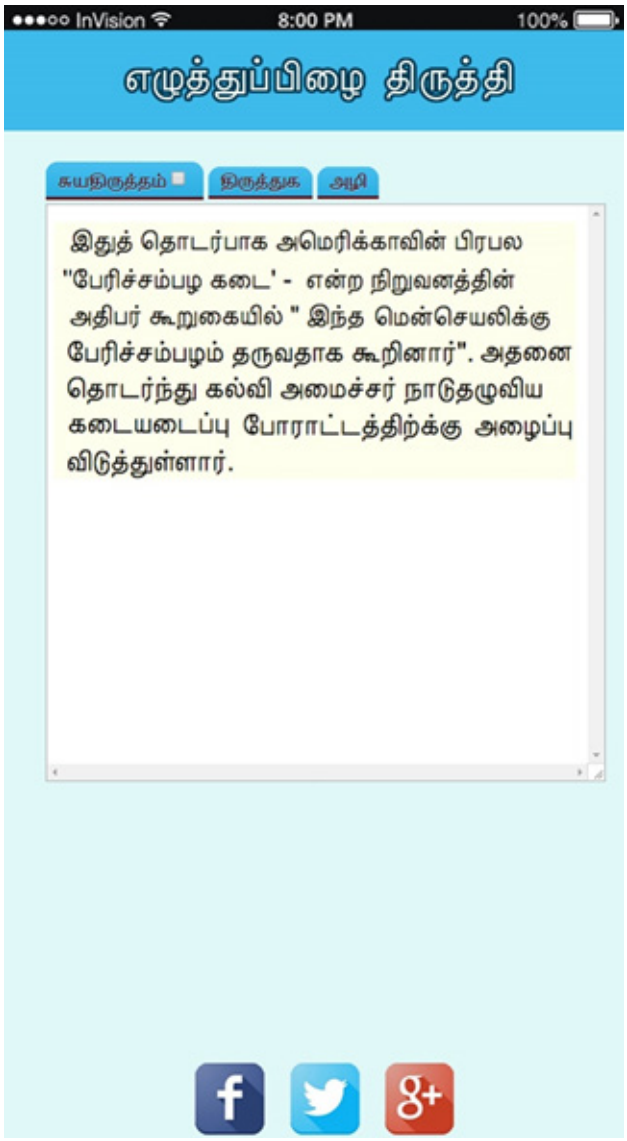

Figure 1. Sample Input Text in Tamil with Spelling Mistakes.

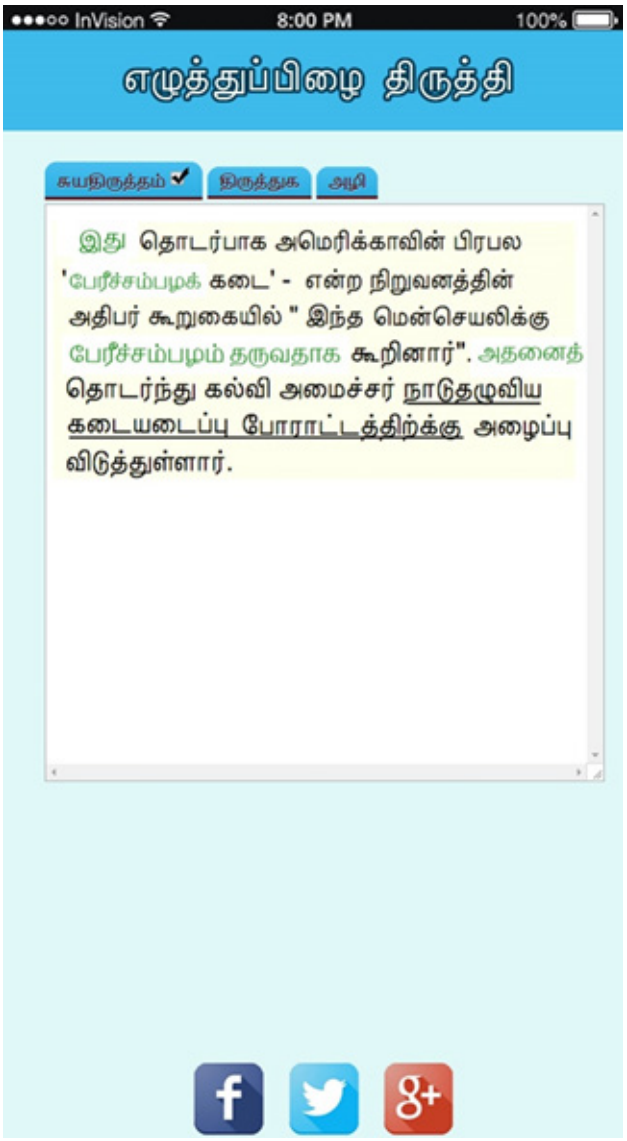

Figure 2. Output Window with Corrected Text on Clicking Autocorrect.

parse the sentences. After checking the words, the rules can execute, and the mistakes can identify. The user can choose auto correction or self-correction or clear options. If the user selects auto correction, then the algorithm will auto correct the words spelling and give the suggestions in various colors. These following colors: green – autocorrected words, red – more than one suggestion, underline – no suggestions but user can justify. Also, this iPhone application will show suggested list of words, if the user can choose the appropriator type their own word. If the user chooses Matru then particular word can change else the user chooses Matru(ye) then the word in all occurrences in entire input can change. The user can delete entire input or output from the window by chooses the erase option[3]. This application cannot identify the other than Tamil words and some of the noun. The results show

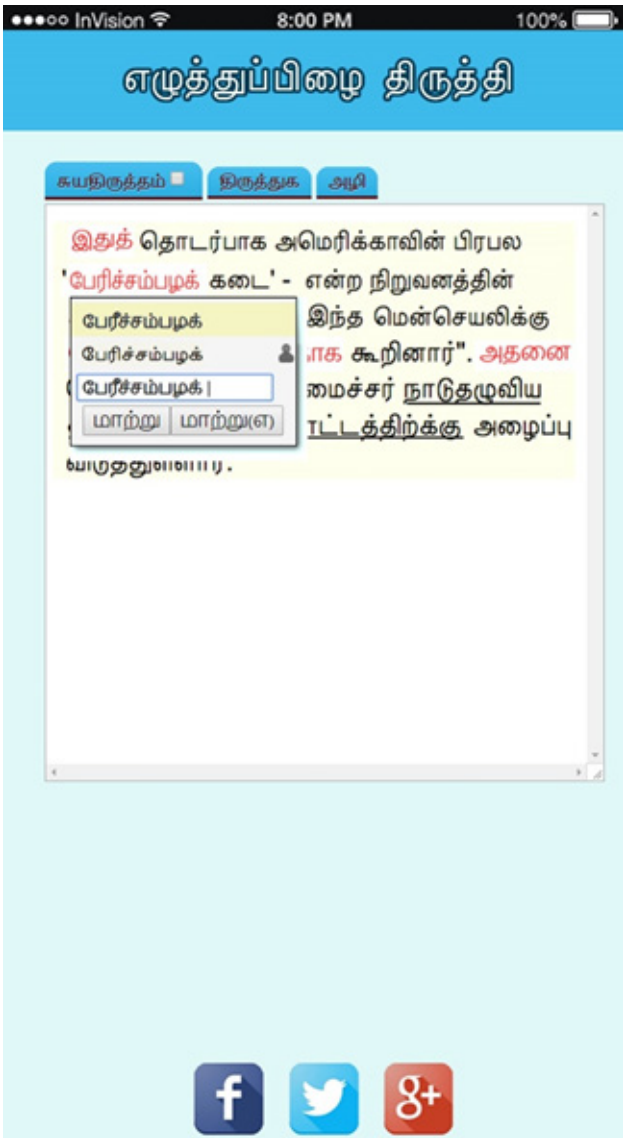

Figure 3. Output Window with Self Corrected Option.

in output window may share by user in Facebook, twitter or Google plus as shown in Figures 1-3.

4. Conclusion

We developed the automated Tamil spell checker app for iPhone mobiles. This analyses the whole input text word by word and combination of two words too. It may further be developed by adding additional words to dictionary. This spell checker can enhance for noun words too.

5. References

1. Tholkapiam. Available from: http://www.tamilvu.org/ta/library-l0100-html-l0100ind-116715. Date accessed: 2017.
2. Nannul. Available from: https://www.projectmadurai.org/pm_etexts/pdf/pm0147.pdf. Date accessed: 2002.
3. Thendral SR, Subhashni V, Karky M, Rajaraman S. Tamil santhi checker app for android phone. International Journal of Innovative Technology and Exploring Engineering. 2019; 8(6):1-3.