

## Tamilnaiya Spellchecker User Manuals

This is a user manuals for the Tamilnaiya Spellchecker developed by Neechalkaran. This desktop based tool has been developed based on online spell checker Vaani, which was released on 2015, and with assistance of Tamil virtual Academy.

### Spellchecker Usage

As a general standard, it identifies the non dictionary words or wrong words and suggest the relevant Tamil word. Basically it suggests and assist on two types of errors. The Grammar based errors like sandhi errors, conjugate errors, consonant errors, homophones etc., are the first type. The errors like typo, unicode errors, non tamil words or dialects are the later. This tool tries to correct these errors and suggest most accurate words.

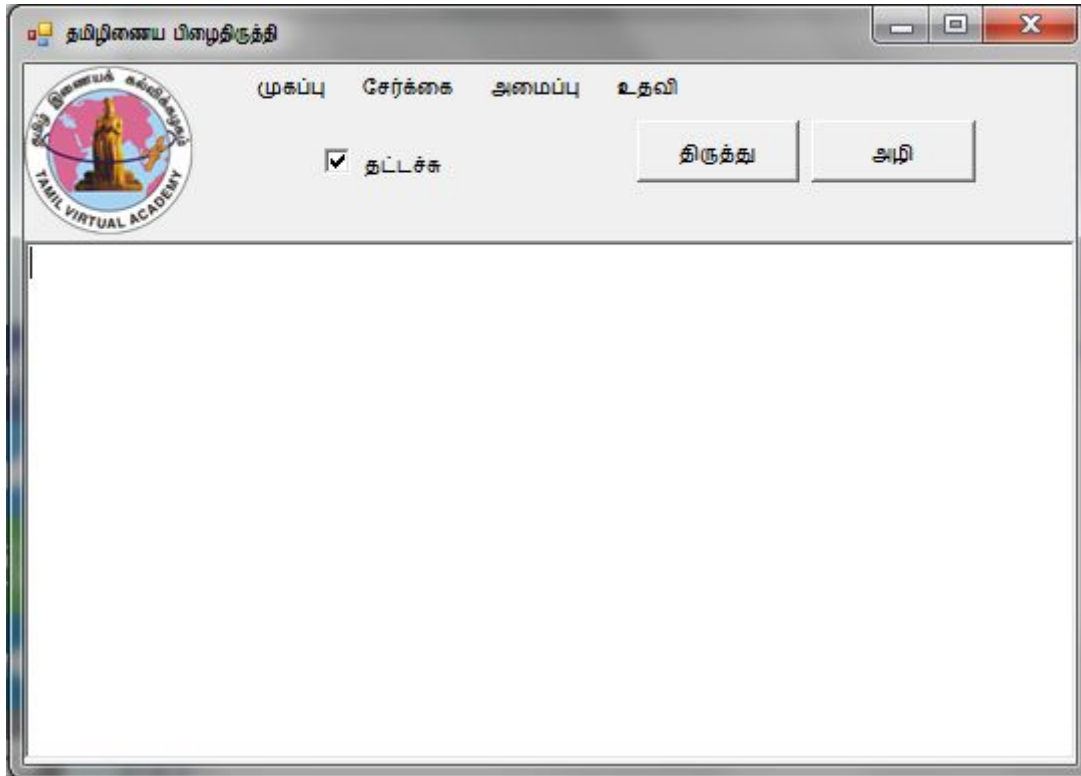
- Sandhi errors are fixed on most efficient way as shown in below. It recognize all mid and suffix sandhi and ensure the content quality  
*Example:* எடுத்துக்கொள்ள -> எடுத்துக்கொள்ள,  
அகட்சி -> அக்கட்சி
- Based on rules for words conjugation, It validates and suggests the correct words as shown in below  
*Example:* நூறுக்கும் - > நூற்றுக்கும்  
It recognizes both valid forms like "கருத்தில் கொண்டு" or "கருத்திற் கொண்டு" however it also identify misplaced consonants while on conjugation "பொருற் பெயர்" get suggestion as "பொருட் பெயர்"
- Homophones like ல-ள-ழ, ன-ந-ண, ர-ற variations are identified more efficiently and suggest the best one.  
*example:* ஒலிந்துவிட்டேன் -> ஒளிந்துவிட்டேன்/ஒழிந்துவிட்டேன், சவறில் -> சவரில்.
- It also suggests relevant Tamil words for few hundred most used Foreign words Mainly english.  
*example:* டாக்டர் -> முனைவர்/மருத்துவர். It also suggests to correct frequent used dialect words. விழுந்திருச்சே -> விழுந்துவிட்டதே, அடிச்சகிட்டு -> அடித்துக்கொண்டு
- Not only visible typo errors however invisible unicode errors are also fixed  
*example:* க் + ஓ represented by க் + எ+ அ கொள்கை -> கொள்கை.
- Other Generic standards in spellchecker. It also underline the words of not found. It can recognize few proper nouns like city name, person name however not all names are identified. So underlined words are not mean as wrong but it not in dictionary. So user can understand accordingly

### Installation

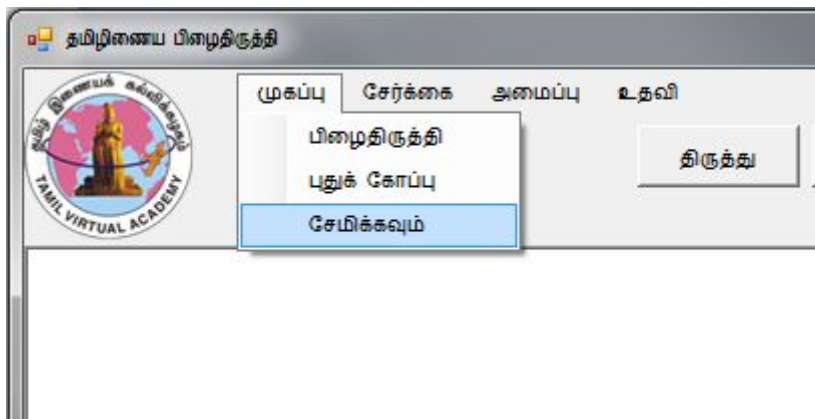
Download the software from Tamil Virtual academy (tamilvu.org) website  
Unzip the file and save it in required folder  
Within the folder, click SpellChecker.exe to open the software.  
It is a portable software and no hard installation required

## Usage

Below is the screenshot of front view



## structure:

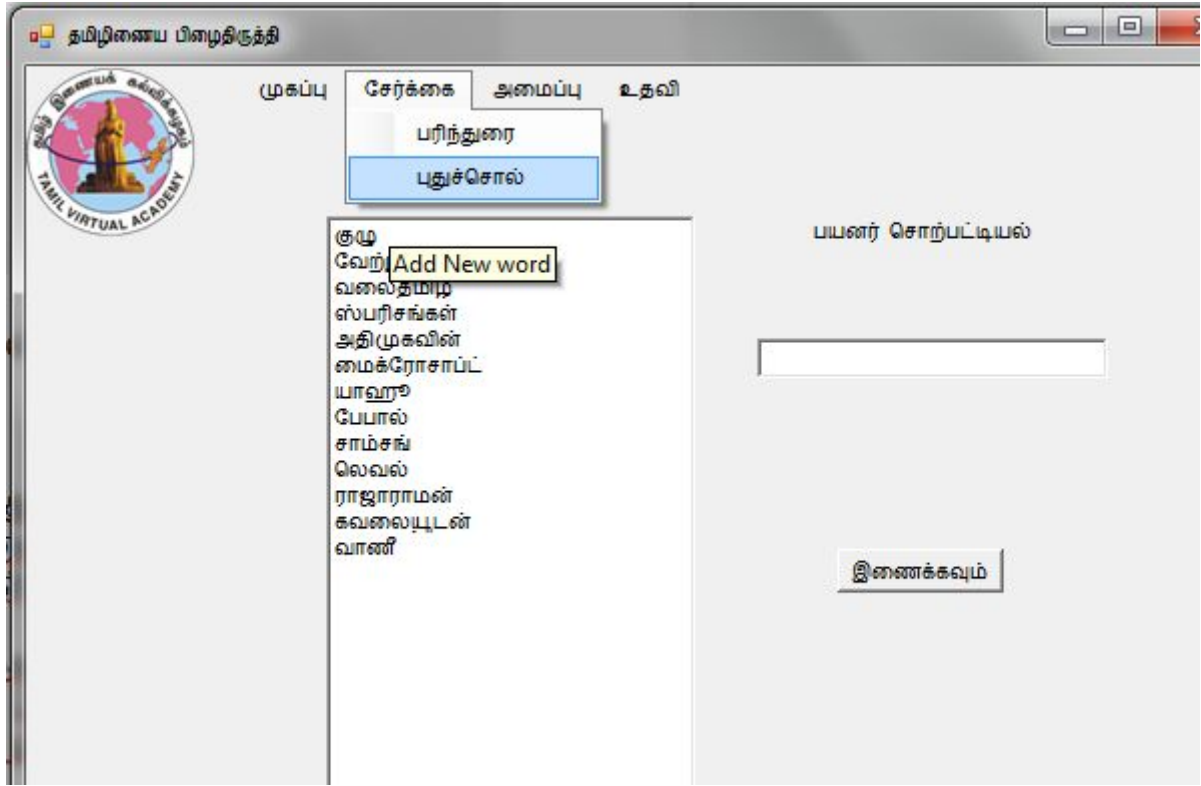


### On Top Menu

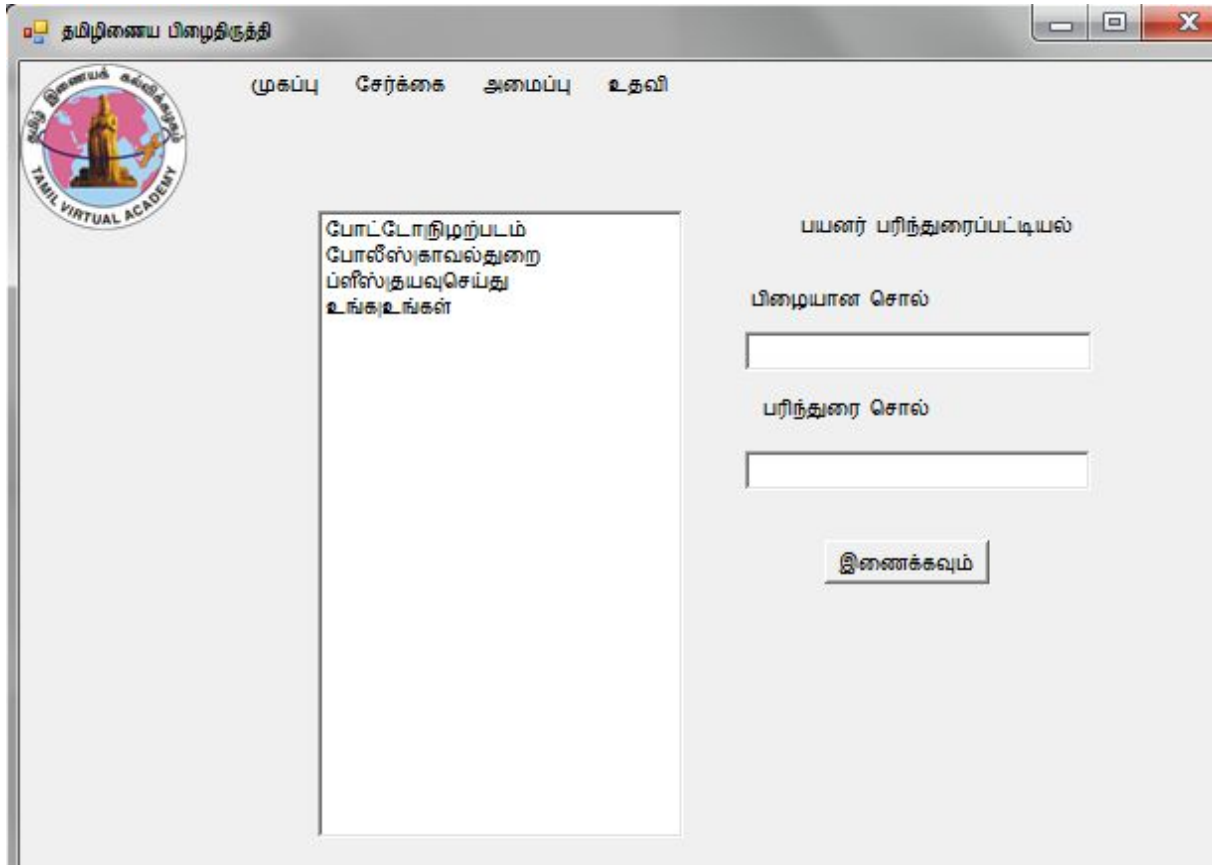
பிழைதிருத்தி: To view the Editor window

புதுக் கோப்பு: To open the new file, click this and select needed text file to be opened

சேமிக்கவும்: Once the activities are done, click thi to save the file again.

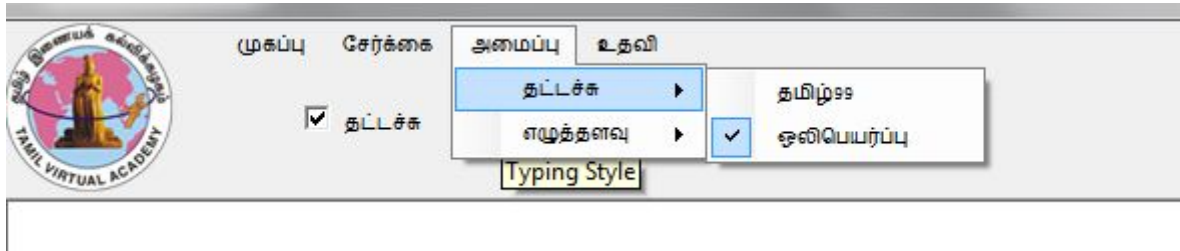


புதுச்சொல்: This is user dictionary. User can add their own dictionary in it. Words can be added directly through above way or even words can be added for live editor windows. Based on these words tool will also suggest the words. List of words can be seen in left side

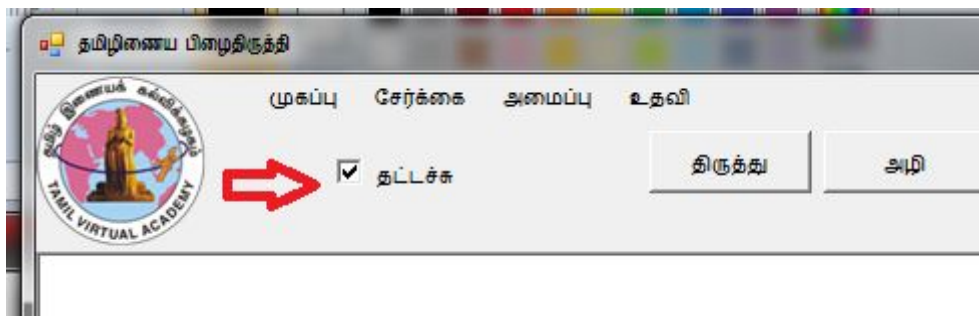


பரிந்துரை: This is for the user preference for custom error. User can add most misspelt words or frequent error words list to this.

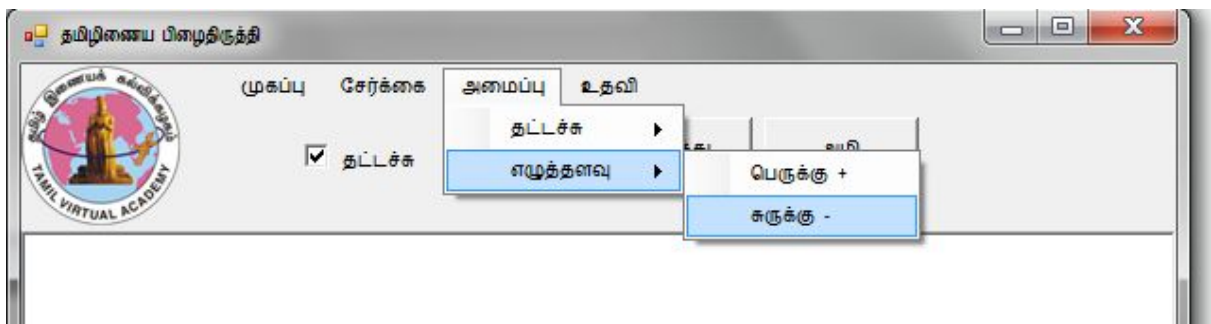
Given wrong words in first textbox and related suggestion needs to be given in second text box. Then click இணைக்கவும் to save it. on left side user can find the preference lists



அமைப்பு: On setting menu, user can choose the typing style. It supports 2 type of typing i.e. phonetic and Tamil99. Phonetic style is the default selection, User can change it accordingly as per their convenience.



Ensure above checkbox is selected to type tamil using available typing keyboard. To type in english or own input keyboard, the user needs to opt out the keyboard selection as shown in image.



To increase or decrease the font size in the editor area.

### How to use

Either type or copy paste or even open a file from elsewhere to this editor window. Once all other typing was completed, click "திருத்து", It may take few seconds to bring the results. In

case to delete the content then click “அழி”. The output results are in three forms black underlined or highlighted in red underlined, green underlined words.

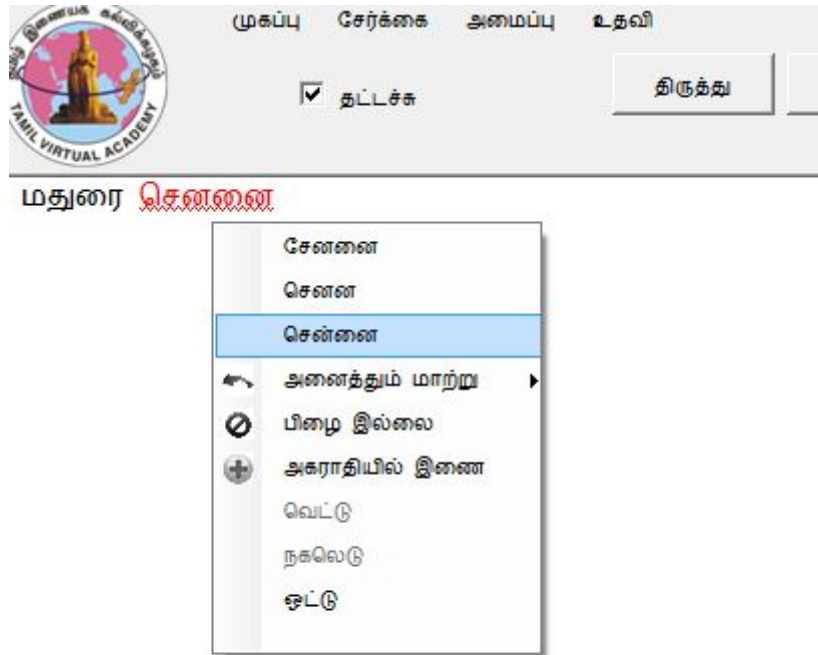
Black highlighted means the word not found in dictionary

Red highlighted means the word not found in dictionary but has suggestion

Green highlighted means the word has been replaced with suggestion

The suggestion dropdown box appears when user click the word, user can choose any of the suggestion or do click "பிழை இல்லை" to ignore it for entire session. If that word needs to be ignored forever, then user needs to click “அகராதியில் இணை” to add it to the custom dictionary

User can also do replace all by using “அனைத்தும் மாற்று” option



## Basic Design

- Unlike other spellchecker, it doesn't do processing only with dictionary words. It use Boyer-Moore algorithm to compare the words and get quick results.
- Based on Nannool rules, all inflected words are derived from every noun and verbs. Which is then used for comparison
- Every noun has around 640 inflect forms and verb has around 1450 inflect forms
- Suggestion words and predicted with Levenshtein distance algorithm from wrong given words

**Users:**

- Tool will be most useful for proof read by Tamil writers and Journalist
- It also helps students and teachers to learn Tamil
- Further development may led to translation and grammar checker

For Latest version of Spellchecker is available at address <http://vaani.neechalkaran.com>  
Suggestions and comments are welcome at [neechalkaran@gmail.com](mailto:neechalkaran@gmail.com)