செம்மொழித் தமிழாய்வு மத்திய நிறுவனம்

மொழித் தொழில்நுட்பப் புலம்

சொல்நிகழ்வெண் கருவி

பயனாளர் கையேடு

http://ctlt.cict.in/

© 2021 செம்மொழித் தமிழாய்வு மத்திய நிறுவனம்
http://ctlt.cict.in/

சொல்நிகழ்வெண் கருவி

1. சொல் நிகழ்வெண்

அகரவரிசை என்பது ஒரு மொழியில் உள்ள எழுத்துக்களை அம்மொழியின் முறைப்படி அடுக்கப்பட்ட எழுத்துக்களின் வரிசை ஆகும். ஏறத்தாழ எல்லா மொழிகளிலுமே அகரம் முதல் எழுத்தாக இருப்பதால் அகரம் தொடங்கி எழுத்துக்கள் வரிசைப்படுத்தப்படுகின்றன. தனித்தனி எழுத்துக்களாகக் கோர்த்துச் சொற்கள் ஆக்கப்படும் மொழிகளுக்கு அகரவரிசை அடிப்படையான ஒன்று. அகரம் தொடங்கி னகரம் வரை சொற்கள் வரிசைப்படுத்தப்படுகின்றன. இவ்வாறு சொற்களை அகரவரிசைப்படி அமைத்து அந்தச் சொற்களின் நிகழ்வெண்ணையும் அளிக்கும் மென்பொருள் சொல்நிகழ்வெண் கருவி எனப்படும்.

http://ctlt.cict.in/

சொல்நிகழ்வெண் கருவி என்பது இயற்கை மொழி ஆய்விற்கான அடிப்படை கருவி. இதனைப் பயன்படுத்தி ஒரு குறிப்பிட்ட நூலின் சொல்நிகழ்வெண் பட்டியலை மட்டும் அல்லாது இதன் இடைமுகத்தில் பல நூல்களை உள்ளீடு செய்து அவற்றின் சொல்நிகழ்வெண்களையும் பெற இயலும். சான்றாக, எட்டுத்தொகையின் எட்டு நூல்களுக்கும் ஒரே நேரத்தில் சொல்நிகழ்வெண்களைப் பெறலாம்.

சொல்நிகழ்வெண் கருவியின் பயன்கள்:

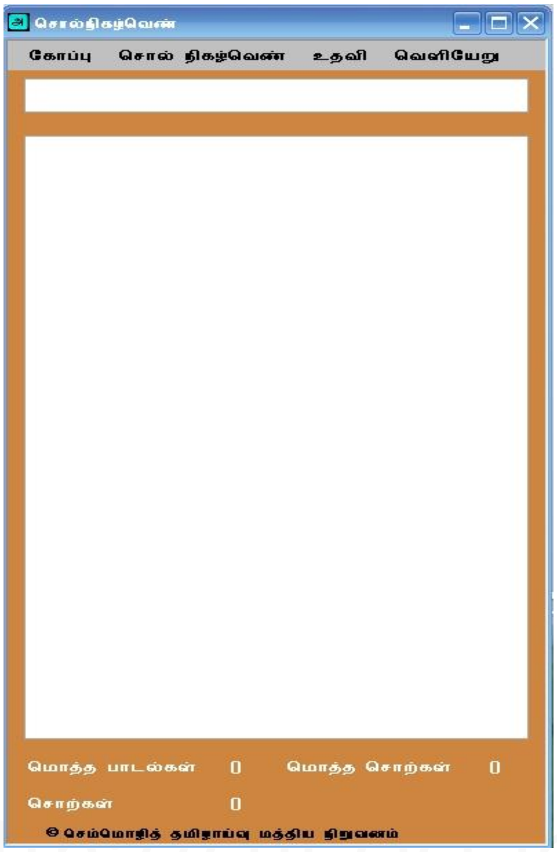
- ஒரே சொல்லின் வெவ்வேறு வகையான பயன்பாடுகளை அறிதல்.
- சொற்பட்டியலின் துணை கொண்டு வேர்சொற்களை இனம் காணுதல்.
- தேடு பொறிகள் உருவாக்கத்தில் அடிப்படை தரவு உருவாக்கத்தில் பயன்படுதல்.
- சொல் வளத்தைச் சிறப்பாகப் பயன்படுத்துதல்.
- மொழிகளில் சிறப்பான தரவுத் தளங்களை உருவாக்குதல்.
- சொல் மற்றும் இலக்கணத் திருத்திகளின் பின்புலத் தரவு உருவாக்த்திற்குப் பயன்படுதல்.
- மொழி உரை ஆய்வு உருவாக்கத்திற்குப் பயன்படுதல்.

http://ctlt.cict.in/

2. சொல்நிகழ்வெண் கருவி - வடிவமைப்பு

சொல்நிகழ்வெண் கருவி பயனாளர்களுக்கு எளிதில் பயன்படும் வகையில் நான்கு (menu bar) கூறுகளாகப் பிரிக்கப்பட்டுள்ளன. அவை

- கோப்பு
- கருவிகள்
- உதவி
- வெளியேறு

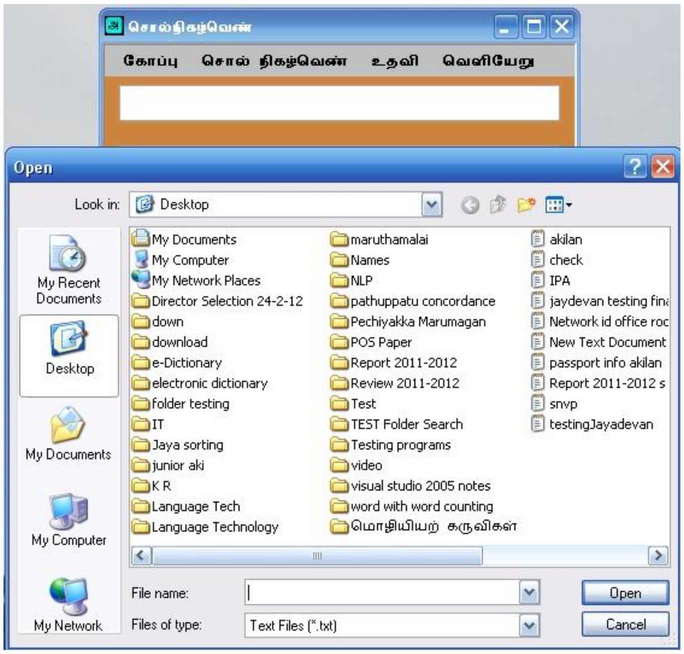


http://ctlt.cict.in/

படம் 1: பயனாளர் இடைமுகம்

- ### கோப்பு

கோப்பைத் தெரிவு செய்க எனும் குமிழ் பனுவலை (Text file) உள்ளீடு செய்வதற்குப் பயன்படுகிறது. ஒருங்குறியில் (UNICODE) அமையப்பெற்ற நூலைப் பனுவல் (*.txt) வடிவில் உள்ளீடு செய்யலாம். சொற்சூழல் கருவியில் ஒன்றிற்கு மேற்பட்ட நூல்களை (multi file support) உள்ளீடு செய்து முடிவுகளைப் பெறும் வசதியும் அளிக்கப்பட்டுள்ளது.



படம் 2: தரவு உள்ளீடு

- **சொல் நிகழ்வெண்**

http://ctlt.cict.in/

கோப்பை உள்ளீடு செய்தவுடன் சொல்நிகழ்வெண் எனும் குமிழைத் தெரிவு செய்தால் அதில் சொல்நிகழ்வெண் வரிசை, முடிவைப்பெற என இரண்டு குமிழ்கள் தோன்றும். இதில் சொல்நிகழ்வெண் எனும் குமிழியைச் சொடுக்கினால் திரையில் சொற்களின் பட்டியல், அச்சொல் பயின்று வரும் நிகழ்வெண்கள் ஆகியன பெறப்படும்.

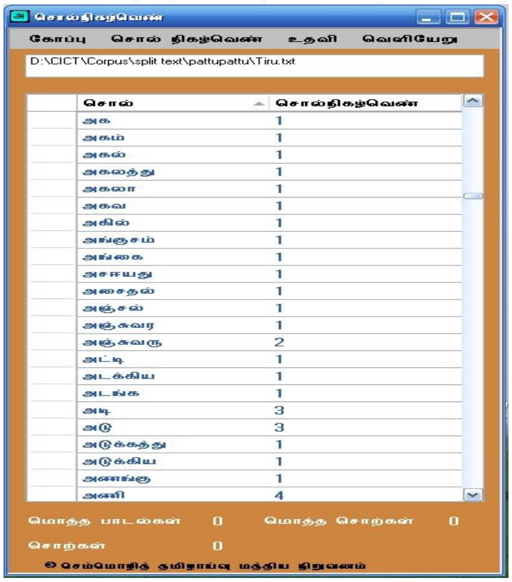
இடைமுகத்தில் உள்ள ‘சொல்’ என்னும் பகுதியினைத் ஒரு முறை சொடுக்கினால் அச்சொற்களை அகர வரிசைப்படுத்தவும், மறு முறை சொடுக்கினால் னகர வரிசைப்படுத்தவும் இயலும்.

‘சொல்நிகழ்வெண்’ என்னும் தலைப்பைச் சொடுக்கினால் உள்ளீடு செய்யப்பட்ட நூலில் எச்சொல் அதிக எண்ணிக்கையில் இடம்பெற்று

இருக்கின்றது என்னும் தகவலைப் பெற இயலும். மறுமுறை சொடுக்கினால் குறைந்த எண்ணிக்கையில் இடம்பெற்று சொற்களின் தகவலைப் பெற இயலும்.

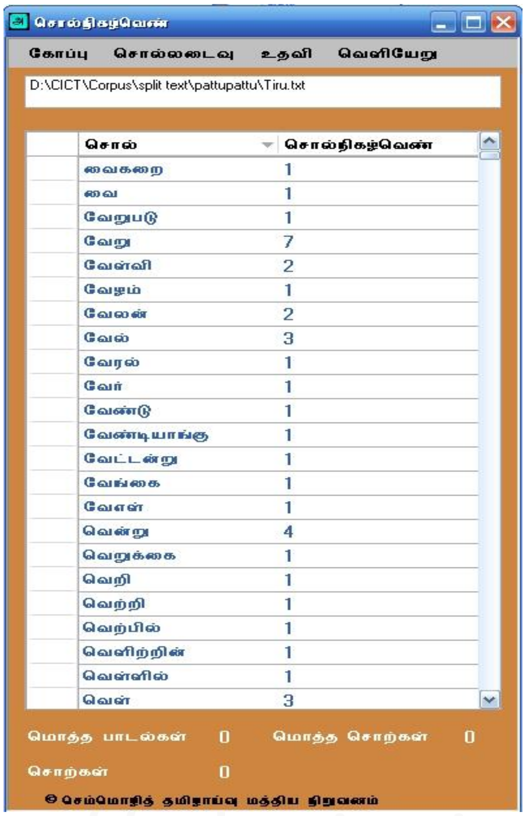
'விழுக்காடு' என்னும் பகுதியில் அகரவரிசை/னகரவரிசைப் படுத்தப்பட்ட அச்சொற்கள் எத்தனை விழுக்காடு என்பதனைப் பெற இயலும். விழுக்காடு பகுதியில் ஏற்கெனவே காணப்பெறும் சொற்களின் விழுக்காட்டினை ஏறு, இறங்குமுகமாக அடைவுபடுத்த இடைமுகத்தின் விழுக்காடு குமிழினை ஒரு முறை சொடுக்க இறங்குமுகமாகவும் மறுமுறை சொடுக்க ஏறுமுகமாகவும் மாறி அமையும்.

3. அகரவரிசை



படம் 3: அகரவரிசை

4. னகரவரிசை

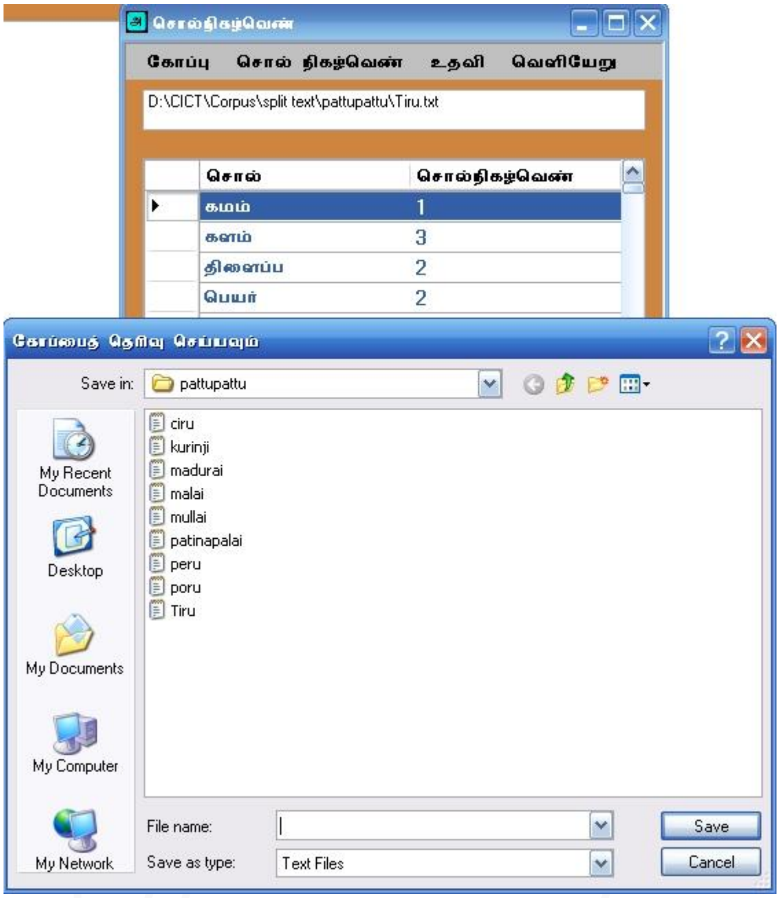


http://ctlt.cict.in/

படம் 4: னகரவரிசை

5. முடிவைப்பெற

சொல்நிகழ்வெண் கருவியில் இருந்து பெறப்பட்ட முடிவைப் பனுவலாகத் (Text file) தரவிறக்கம் செய்து கொள்ளும் வாய்ப்பும் அளிக்கப்பட்டுள்ளது. 'முடிவைப்பெற' எனும் குமிழியைத் தெரிவுசெய்தால் 'கோப்பைத் தெரிவு செய்யவும்' எனும் தலைப்பிட்ட இடைமுகம் தோன்றும். இதில் கோப்பின் பெயரைத் தட்டச்சுச் செய்து சேமி (save) எனும் குமிழினைத் தெரிவு செய்தால் சொற்சூழல்கருவியின் மூலம் பெறப்பட்ட முடிவை பனுவலாகப் (Text file) பெறலாம்.



படம் 5: முடிவைப்பெற

6. நீக்கு

'நீக்கு' எனும் குமிழியைத் தெரிவு செய்வதன் மூலம் சொல்நிகழ்வெண் கருவியில் பயன்படுத்தப் பட்டுள்ள அனைத்துப் பகுதிகளையும் நீக்கலாம். புதிய நூலின் பயனைப் பெறவேண்டுமெனில் முன் பயன்படுத்தப்பட்ட பகுதியினை நீக்குதல் இன்றியமையாதது.

7. வெளியேறு

'வெளியேறு' எனும் குமிழியைத் தெரிவு செய்வதன் வாயிலாகச் சொல்நிகழ்வெண் கருவியில் இருந்து வெளியேறலாம்.

சொல்நிகழ்வெண் கருவி

பதிப்பு 1.0

காப்புரிமை **செம்மொழித் தமிழாய்வு மத்திய நிறுவனம்**

உருவாக்கம் மொழித் தொழில்நுட்பப் புலம்

தொடர்புக்கு செம்மொழித் தமிழாய்வு மத்திய நிறுவனம்
செம்மொழிச் சாலை
பெரும்பாக்கம்
சென்னை - 600100
தொலைபேசி : +91 - 44 22540125

மின்னஞ்சல் : dlt@cict.in

இணையம் : https://www.cict.in/

https://twitter.com/CICTOfficial

https://www.instagram.com/cict_chennai/

https://www.facebook.com/chennaicict/

https://www.youtube.com/user/cictarchives

http://clt.cict.in/