



செம்மொழித் தமிழாய்வு மத்திய நிறுவனம்
மொழித் தொழில்நுட்பப் புலம்

சொல்லடைவி

<http://ctlt.cict.in/>

பயனாளர் கையேடு

© 2021 செம்மொழித் தமிழாய்வு மத்திய நிறுவனம்

<http://ctlt.cict.in/>

தமிழ்ச் சொல்லடைவி

1. சொல்லடைவி

சொல்லடைவு என்பது ஒரு நூலில் அல்லது பதிப்பில் பயன்படுத்தப்பட்டுள்ள சொற்களின் அகர வரிசைப்படுத்தப்பட்ட பட்டியலாகும். இப்பட்டியலில் சொற்களுடன் அவை இடம்பெறும் தொடர் எண் மற்றும் அடி குறிப்பிடப்பட்டிருக்கும். அதுபோல, சொல்லடைவி என்பது கொடுக்கப்பட்ட ஒரு நூலின் சொல் அடைவை உருவாக்கித் தரக்கூடிய கணினி நிரல். இப்பணியைச் செய்துதரும் மென்பொருள் சொல்லடைவி எனப்படும்.

சொல்லடைவியின் பயன்கள்:

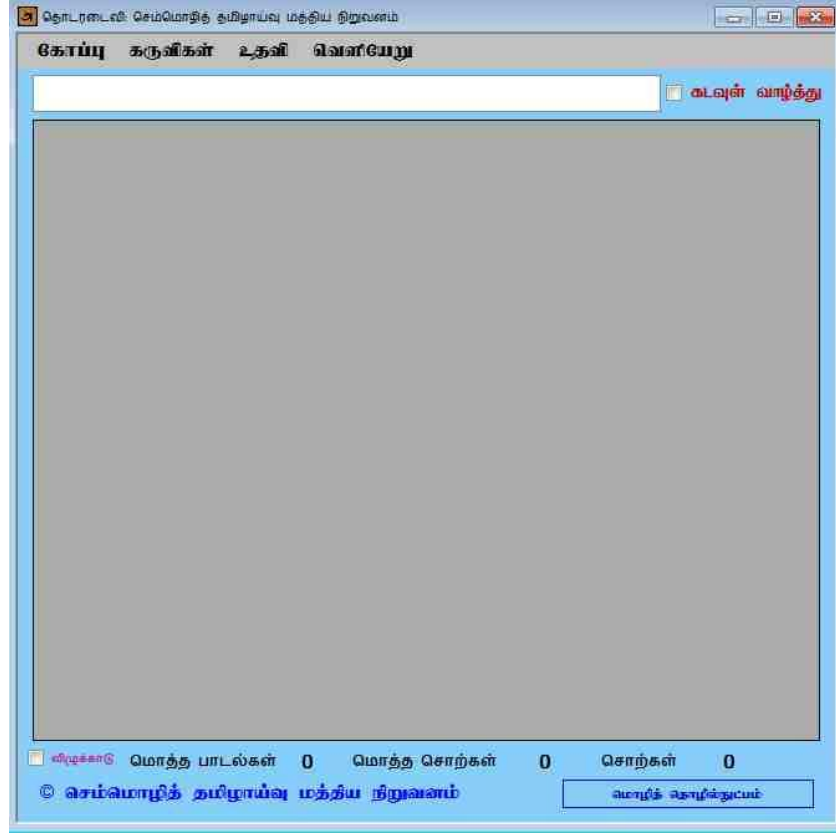
- ஒரு சொல் சங்க இலக்கியங்களில் எத்தனை முறை வருகின்றது என்பதைச் சொல்லடைவு மூலம் அறியலாம்.
- மொழியியல் விதிகள், கோட்பாடுகள் சரியானவைதானா என்பதைச் சொல்லடைவு மூலம் அறியலாம்.
- வரலாற்று முறை இலக்கணம், வரலாற்று முறை அகராதி தயாரிப்பதற்குச் சொல்லடைவு பயன்படுகிறது.
- சீர், உருபன், சொல், தொடர் முதலியவற்றைப் பிரிக்கப் பயன்படும் கருவிகளை உருவாக்குவதற்கும் சொல்லடைவு பயன்படுகின்றது.
- வேர்ச் சொல்லாய்வுகள், பல்வேறு அகராதிகள், பொருட்களஞ்சியங்கள் தயாரிப்பதற்கும் சொல்லடைவு பயன்படுகின்றது.
- கல்வி நுட்பவியல், மொழி நுட்பவியல் துறைகளுக்கும், மின் அகராதி, சொல் திருத்தி, இலக்கணத் திருத்தி, பேச்சு மொழியிலிருந்து எழுத்து மொழியை உருவாக்கும் கருவி. எழுத்து மொழியிலிருந்து பேச்சு மொழியை உருவாக்கும் கருவி,

மொழிபெயர்ப்புக் கருவி போன்ற மென்பொருள் உருவாக்கத்திற்கும் பயன்படுகிறது.

■ அமைப்பு

சொல்லடைவி பயனாளர்களுக்கு எளிதில் பயன்படும் வகையில் நான்கு (menu bar) கூறுகளாகப் பிரிக்கப்பட்டுள்ளது. அவை

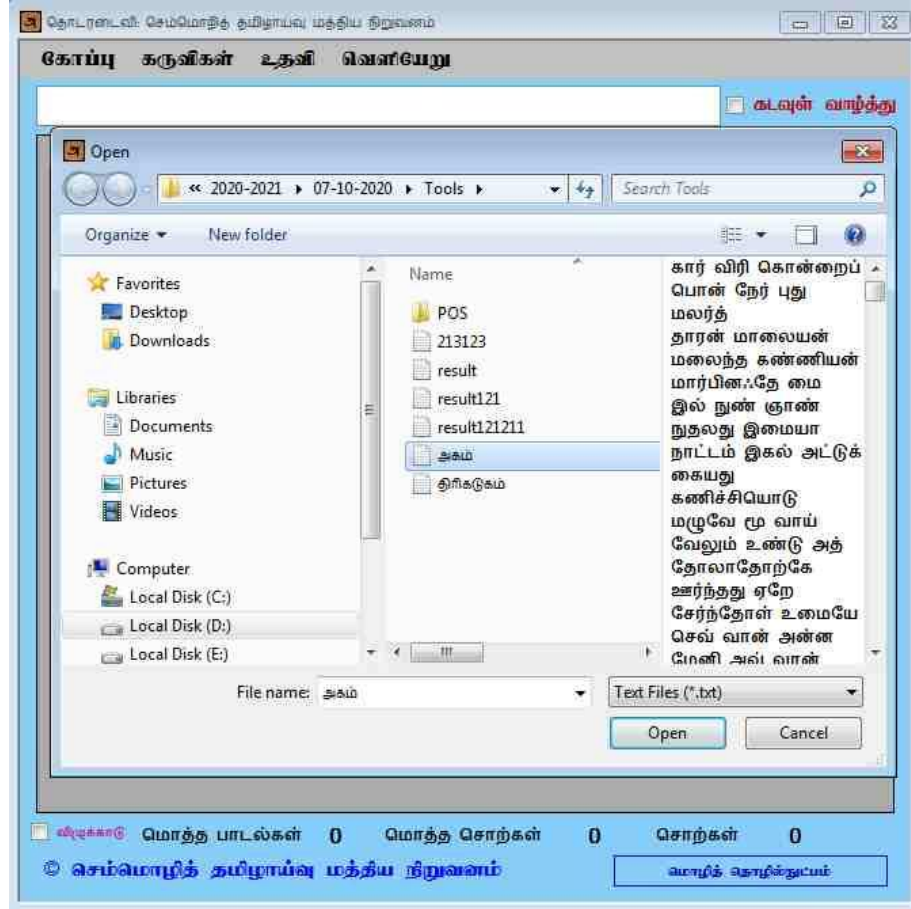
- கோப்பு
- கருவிகள்
- உதவி
- வெளியேறு



படம் 1. சொல்லடைவி

■ கோப்பு

கோப்பைத் தெரிவு செய்க எனும் குமிழ் பனுவலை (Text file) உள்ளீடு செய்வதற்கு பயன்படுகிறது. ஒருங்குறியில் (UNICODE) அமையப்பெற்ற நூலை பனுவல் (*.txt) வடிவில் உள்ளீடு செய்யலாம். இக் கருவியில் ஒன்றிக்கு மேற்பட்ட நூல்களை (multi file support) உள்ளீடு செய்து முடிவுகளை பெறும் வசதியும் அளிக்கப்பட்டுள்ளது.



படம் 2. தரவு உள்ளீடு

கருவிகள்

சொல்லடைவு என்பது நான்கு உட்பிரிவுகளாகப் பிரிக்கப்பட்டுள்ளது.

- ❖ சொல்லடைவு
- ❖ சொல்லடைவைப் பெற
- ❖ தேடு
- ❖ கணக்கீடு

❖ சொல்லடைவு

கோப்புத் தெரிவு செய்யப்பட்டவுடன் 'சொல்லடைவு' எனும் குமிழியைத் தெரிவு செய்தால் உள்ளீடு செய்யப்பட்ட நூலுக்கான சொல் அடைவைப் பெறலாம். இதில் சொல், விழுக்காடு, சொல்நிகழ்வெண், சொல் வருகை ஆகிய நான்கு பகுதிகளைப் பெறலாம்.

<http://ctlt.cict.in/>

❖ சொல்

சொல் என்பது உள்ளீடு செய்யப்பட்ட கோப்பில் உள்ள சொற்களின் வரிசைப்படுத்தப்பட்ட பட்டியலாகும்.

❖ அகரவரிசை/எகரவரிசை

இடைமுகத்தில் உள்ள 'சொல்' என்னும் பகுதியினை ஒரு முறை சொடுக்கினால் அச்சொற்களை அகர வரிசைப்படுத்தவும், மறு முறை சொடுக்கினால் எகர வரிசைப்படுத்தவும் இயலும்.

❖ விழுக்காடு

‘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.

❖ சொல்நிகழ்வெண்

சொல்நிகழ்வெண் என்பது உள்ளீடு செய்யப்பட்ட கோப்பில் உள்ள மொத்த சொற்களில் குறிப்பிட்ட அந்தச் சொல் பயின்று வந்துள்ள சொற்களின் எண்ணிக்கையே.

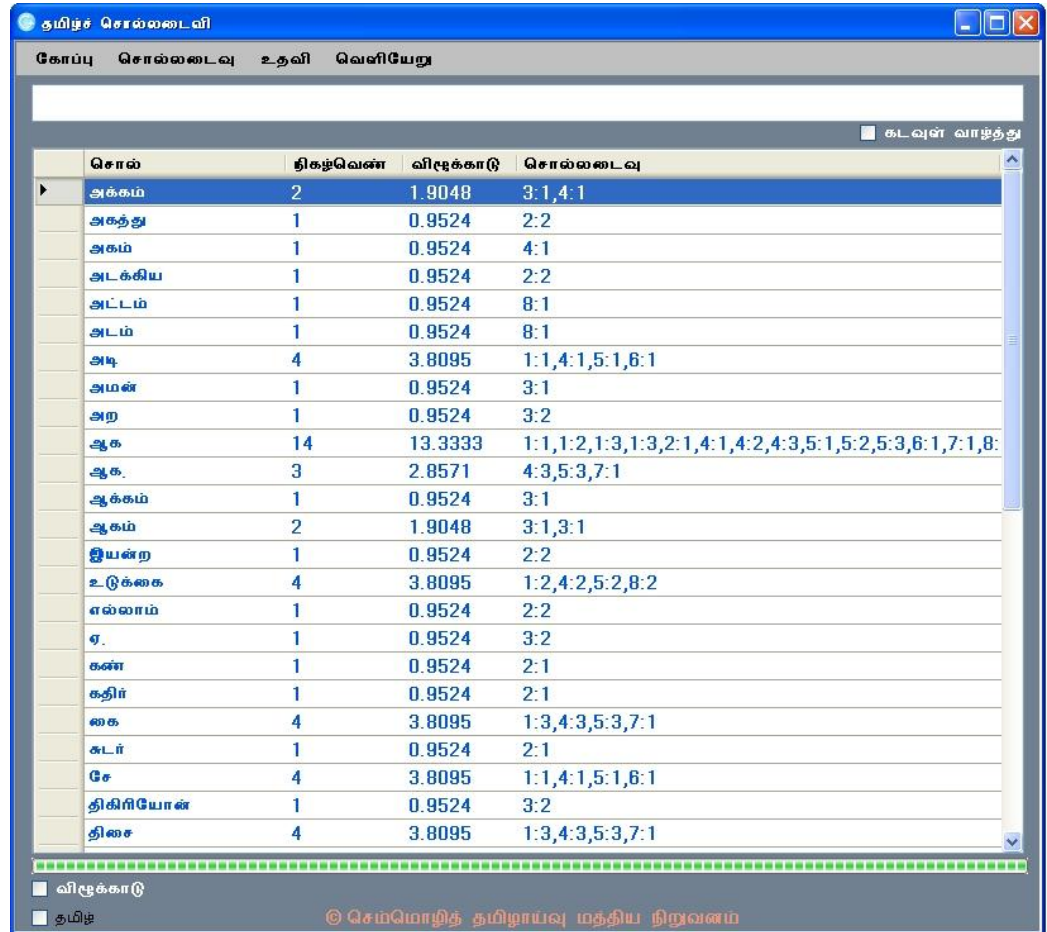
❖ கடவுள் வாழ்த்து

‘கடவுள் வாழ்த்து’ எனும் விருப்பதேர்வைத் தெரிவு செய்து சொல்லடைவைச் சுழியத்தில் இருந்தும் அல்லது ஒன்றில் இருந்தும் தொடங்குமாறு பயனாளர் தம் விருப்பதிற்கேற்ப அமைத்துக் கொள்ளலாம். சங்க இலக்கியங்களைப் பதிப்பிக்கும் போது கடவுள் வாழ்த்தை உட்படுத்தி எண்ணிட்டும், கடவுள் வாழ்த்தினை விடுத்து எண்ணிட்டுப் பதிப்பிக்கும் மரபு காணப் பெறுகின்றது. ஆகவே இச்சொல்லடைவுக் கருவியினைப் பயன்படுத்தும் போது இத்தகைய குழப்பம் வராமல் இருக்க மேற்குறிப்பிட்டவாறு இரு முறையில் எண்ணிட்டுத் தரப்பெற்றுள்ளது. இது இதன் தனிச் சிறப்பு.

❖ சொல் வருகை

சொல் வருகை என்பது உள்ளீடு செய்யப்பட்ட கோப்பில் உள்ள மொத்த சொற்களில் குறிப்பிட்ட அந்தச் சொல் இடம்பெற்ற பாடல் எண், அடி எண் ஆகியவற்றைக் கொண்டது. இந்தக் குறிப்பிட்ட எண்ணின் வரிசை ஒன்றில் இருந்து தொடங்கும் வசதி செய்யப்பட்டுள்ளது.

சொல்லடைவுக் கருவியில் நம் முடிவுகளுக்குத் தொடர்பில்லாத தரவுகள் இடம்பெறின், இத்தேவையற்ற முடிவுகளை நீக்க விசைப்பலகையின் 'Delete' குமிழினைப் பயன்படுத்தலாம்.



The screenshot shows a window titled 'தமிழ்ச் சொல்லடைவி' (Tamil Word Frequency Software). The window has a menu bar with 'கோப்பு' (File), 'சொல்லடைவு' (Word Frequency), 'உதவி' (Help), and 'வெளியேறு' (Exit). Below the menu bar is a toolbar with a 'கடவுள் வாழ்த்து' (Praise God) button. The main area contains a table with the following columns: 'சொல்' (Word), 'திகழ்வெண்' (Frequency), 'விலக்காடு' (Exclusion Code), and 'சொல்லடைவு' (Word Frequency). The table lists various Tamil words and their frequencies, along with exclusion codes and word frequency details. At the bottom of the window, there are checkboxes for 'விலக்காடு' (Exclusion Code) and 'தமிழ்' (Tamil), and a copyright notice: '© செம்மொழித் தமிழாய்வு மத்திய நிறுவனம்'.

சொல்	திகழ்வெண்	விலக்காடு	சொல்லடைவு
அக்கம்	2	1.9048	3:1,4:1
அகத்து	1	0.9524	2:2
அகம்	1	0.9524	4:1
அடக்கிய	1	0.9524	2:2
அட்டம்	1	0.9524	8:1
அடம்	1	0.9524	8:1
அடி	4	3.8095	1:1,4:1,5:1,6:1
அமன்	1	0.9524	3:1
அற	1	0.9524	3:2
ஆக	14	13.3333	1:1,1:2,1:3,1:3,2:1,4:1,4:2,4:3,5:1,5:2,5:3,6:1,7:1,8:
ஆக.	3	2.8571	4:3,5:3,7:1
ஆக்கம்	1	0.9524	3:1
ஆகம்	2	1.9048	3:1,3:1
ஐயன்ற	1	0.9524	2:2
உடுக்கை	4	3.8095	1:2,4:2,5:2,8:2
எல்லாம்	1	0.9524	2:2
ஏ.	1	0.9524	3:2
கண்	1	0.9524	2:1
கதிர்	1	0.9524	2:1
கை	4	3.8095	1:3,4:3,5:3,7:1
கடர்	1	0.9524	2:1
சே	4	3.8095	1:1,4:1,5:1,6:1
திகிரியோன்	1	0.9524	3:2
திசை	4	3.8095	1:3,4:3,5:3,7:1

படம் 3. சொல்லடைவு

■ முடிவைப்பெற

சொல்லடைவிக் கருவியில் இருந்து பெறப்பட்ட முடிவைப் பனுவலாகத் (Text file) தரவிறக்கம் செய்து கொள்ளும் வாய்ப்பு அளிக்கப்பட்டுள்ளது. ‘முடிவைப்பெற’ எனும் குமிழியைத் தெரிவுசெய்தால் ‘கோப்பைத் தெரிவு செய்யவும்’ எனும் தலைப்பிட்ட சன்னல் திரை தோன்றும். இதில் கோப்பின் பெயரைத் தட்டச்சுச் செய்து சேமி (save) எனும் குமிழினைத் தெரிவு செய்தால் சொல்லடைவி கருவியின் மூலம் பெறப்பட்ட முடிவை பனுவலாகப் (Text file) பெறலாம். இதில் சொல், சொல்நிகழ்வெண், சொல்லடைவு ஆகிய மூன்று பகுதிகளை பெற இயலும்.

கூடுதலாக இக்கருவியில் சில ஆய்வுகளுக்கு ‘விழுக்காடு’ எனும் முடிவு தேவைப்படும் நிலையில் ‘விழுக்காடு’ விருப்பத்தேர்வு அளிக்கப்பெற்றுள்ளது. அதனைத் தெரிவு செய்து அனைத்துப் பகுதிகளை பனுவலாகப் (Text file) பெறலாம்.

இதன் முடிவுகளைப் பனுவலாகப் பெறும்போது (Text file) நூலின் மொத்த பாடல் எண்ணிக்கையும், மொத்த சொற்களின் எண்ணிக்கையும், ஒருமுறை மட்டுமே இடம்பெற்றுள்ள சொற்களின் எண்ணிக்கையும் பெறலாம். சொல்லடைவி கருவி தரவினை உருவாக்க எடுத்துக் கொண்ட கால அளவு, அதன் பயனை அளிக்க எடுத்துக்கொண்ட கால அளவு, இவ்விரண்டும் செயல்படும் இடைப்பட்ட கால அளவு ஆகிய தகவல்களையும் தருவதாக இக்கருவி அமைகின்றது.

❖ கணக்கீடு

சொல்லடைவியைப் பயன்படுத்தித் சொல்லடைவைப் பெற்ற பின் 'கணக்கீடு' எனும் குமிழியைத் தெரிவு செய்தால் உள்ளீடு செய்யப்பட்ட நூலின் மொத்த பாடல் எண்ணிக்கையும், மொத்த சொற்களின் எண்ணிக்கையும், ஒருமுறை மட்டுமே இடம்பெற்றுள்ள சொற்களின் எண்ணிக்கையும் பெறலாம்.

❖ தேடு

சொல்லடைவியில் இருந்து பெறப்பட்ட முடிவில் இருந்து ஒரு குறிப்பிட்ட சொல்லை மட்டும் தேடி அதன் முடிவைப்பெறும் வசதி அளிக்கப்பட்டுள்ளது. இதில் தேடு எனும் குமிழைத் தெரிவு செய்தால் ஒரு திரை தோன்றும். இந்தத் திரையில் சொல், தேடு எனும் இரண்டு குமிழ்கள் இடம் பெற்றுள்ளன. சொல் என்னும் குமிழின் அருகில் கொடுக்கபெற்றிருக்கும் உரைப்பெட்டியில் தேடும் சொல்லினை உள்ளீடு செய்தால், அச்சொல், பயின்று வரும் நிகழ்வெண்கள், விழுக்காடு, சொல்வருகை ஆகியன பெறப்படும். இந்தத் திரையிலும் தேடுதலின் நிறைவாகப் பெறும் முடிவினையும் தரவிறக்கம் செய்ய இயலும். 'வெளியேறு' என்னும் குமிழினைத் தெரிவு செய்தால் 'தேடு' என்னும் பகுதியில் இருந்து மட்டும் வெளியேறலாம்

தொடர்புக்கு

தொடர்புக்கு எனும் குமிழினைச் சொடுக்கினால் உருவாக்கம் எனும் குமிழ்த் தோன்றும். இதனைத் தெரிவு செய்தால் இந்தக் கருவியை உருவாக்கிய குழு மற்றும் நிறுவனத் தொடர்பு முகவரி, மின்னஞ்சல், தொலைபேசி எண்களைப் பெறலாம். இந்த மின்னஞ்சல் முகவரியில் பயனாளர்கள் தங்கள் கருத்துக்களைப் பதிவு செய்யலாம்.

வெளியேறு

வெளியேறு எனும் பட்டியலில் நீக்கு, வெளியேறு என இரண்டு குமிழிகள் அளிக்கப்பட்டுள்ளன. இதில் 'நீக்கு' எனும் குமிழியைத் தெரிவு செய்வதன் மூலம் சொல்லடைவியின் அனைத்து முடிவுகளையும் நீக்க இயலும். சொல்லடைவியில் இருந்து வெளியேற 'வெளியேறு' எனும் குமிழியைத் தெரிவு செய்வதன் மூலம் வெளியேறலாம்.

<http://ctlt.cict.in/>

சொல்லடைவி

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>