

=====

Language in India www.languageinindia.com ISSN 1930-2940 Vol. 23:3 March 2023

=====

கணினித் தொழிநுட்பம் அடிப்படையில் சொல்வளம்சார்
மொழிக்கருவிகள் உருவாக்குதல்

(Developing Lexical Resourceful Aids Based on Computational
Technology)

இராசேந்திரன் சங்கரவேலாயுதன்

அமிர்தா பல்கலைக்கழகம்

கோயம்புத்தூர்

rajushush@gmail.com

Rajendran Sankaravelayuthan

Retired Professor, Tamil University, Thanjavur

Professor of Linguistics

Centre for Excellence in Computational Engineering and Networking (CEN)

Amrita University

Coimbatore 641112

Mobile: 0-9486332155

=====

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil
Textbook)

Abstract of the booklet in English

The booklet is on “Developing lexical resourceful aids based on computational technology”. The booklet has five units excluding a small introduction. The first unit is on “Developing Indexing, concordance and compendium”. The second unit is on “Dictionary and dictionary making”. The third unit is on “Thesaurus and Thesaurus making”. The unit chapter is on “Developing Visual Onto-Thesaurus for Tamil”. The sixth unit is on “WordNet and developing WordNet”

ஆசிரியர் உரை

அகராதி, சொற்களஞ்சியம், சொல்வலை என்பன பற்றி நான் எழுதிய கட்டுரைகளை ஒன்றிணைத்து இன்றைய அளவில் மேம்படுத்தி மாற்றங்கள் செய்து 'கணிணித் தொழிநுட்பம் அடிப்படையில் சொல்வளம்சார் மொழிக்கருவிகள் உருவாக்குதல்' என்ற தலைப்பில் ஒரு சிற்றேடாக நான் உங்கள் முன் சமர்ப்பிக்கின்றேன். இதில் அறிமுகம் தவிர்த்து ஐந்து இயல்கள் உள்ளன. பேராசிரியர் ராஜிவ் சங்கல் அவர்கள் ஒருதடவை சொன்னார் உங்கள் மொழி நிலைத்து நிற்கவேண்டுமானால் ஒரே வழி அம்மொழி தொடர்பாக ஏராளமாக எழுதி இணையத்தில் பதிவிடுங்கள். உங்கள் மொழி இணையத்தில் வாழவும் நீங்கள் இறவாது இணையத்தில் வாழவும் ஒரே வழி அதுதான் என்றார். அது எனக்குச் சரியாகத் தோன்றியது. எனவே தமிழ் மொழியியியல் பற்றி கட்டுரைகளும் நூல்களும் எழுதி பேராசிரியர் எம்.எஸ். திருமலை அவர்களின் Language in India என்ற மின் திங்கள் இதழில் தொடர்ந்து வெளியிட்டு வருகின்றேன். பலர் என் மின் நூல்களைப் பயன்படுத்தி என் முயற்சிகளுக்கு நன்றி தெரிவித்து வருகின்றனர். அவர்கள் தந்த ஊக்கம் தான் மேலும் மேலும் என்னை எழுத் தூண்டுகிறது. அவர்களுக்கு எனது நன்றி.

அன்புடன்

இராசேந்திரன் சங்கரவேலாயுதன்

பொருளடக்கம்

இயல்	தலைப்பு	பக்கம்
	அறிமுகம்	5
1	அகரவரிசை அட்டவணை, சொல்லடைவு, சொற்பொருளடைவு உருவாக்கம்	6
2	அகராதியும் அகராதி உருவாக்கமும்	11
3	சொற்களஞ்சியமும் சொற்களஞ்சிய உருவாக்கமும்	73
4	தமிழ் காட்சி மூலப்பொருண்மையியல் சொற்களஞ்சிய உருவாக்கம்	117
5	சொல்வலையும் சொல்வலை உருவாக்கமும்	137
	நோக்கீட்டு நுல்களும் கட்டுரைகளும்	158

அறிமுகம்

கணியின் செயல்பாட்டுக்கு அடிப்படையாக மொழியில் அமைந்துவந்தது. உள்ளிடப்படும் எதையும் 1, 0 எண்களால் அறிந்துகொள்வதுதான் கணிப்பொறியின் அடிப்படை அறிவு. மனித மூளை 0, 1 என்ற அடிப்படை எண்களை வைத்துக்கொண்டு எந்திர மொழிமூலம் சிக்கலான கணிப்பொறியை உருவாக்கியது. எழுத்துக்களும் எண்களும் தனித்தன்மையான குறியீடுகளும் (Special Symbols) 0, 1 என்பவைகளால்தான் உருப்படுத்தம் செய்யப்படுகின்றன. உயர்நிலை மொழிகள், பொருள்கோளிகள் (Interpreters) ஒன்றுகூட்டிகள் (Assemblers) என்பன மொழியியலால் பயன்பெற்றன. இப்பொழுது மொழியியலார் கணிப்பொறியை/கணினியை அவர்களின் கடுமையான செயல்பாட்டை செய்யும் வன்மையான கருவியாகப் பயன்படுத்தி வருகின்றனர்.

சொல்லடைவு, சொல்லாட்சி அகராதி, பொருளடைவு, அகராதி, சொற்களஞ்சியம், சொல்லவலை என்பனவற்றை சொல்வளம் சார் மொழிக்கருவிகளாகப் பட்டியலிடலாம். இச்சொல்வளங்களை, குறிப்பாக அகராதி, சொற்களஞ்சியம் போன்ற சொல்வளங்களை உருவாவாக்குவதற்காகன தேவைகளும் வழிமுறைகளும் இந்நூலில் விவரிக்கப்படும்.

1. அகரவரிசை அட்டவணை, சொல்லடைவு, சொற்பொருளடைவு உருவாக்கம்

தொடக்க காலங்களில் அகரவரிசை அட்டவணை, சொல்லடைவு, சொற்பொருளடைவு என்பன உருவாக்குதல் மனித முயற்சியால் மேற்கொள்ளப்பட்டு வந்தது. எடுத்துக்காட்டாக தமிழில் சங்க இலக்கிய நூல்களான புறநானூறு, கலித்தொகை போன்ற நூல்களில் உள்ள சொற்களை அட்டவணைப்படுத்த இந்நூல்களில் உள்ள பாக்கள் மனித முயற்சியால் சொற்களாகப் பிரிக்கப்பட்டு அவற்றை அட்டைகளில் அவை நூலில் வரும் இடம், வரும் தொடர், அர்த்தம் ஆகியன எழுதப்பட்டு அகரவரிசைப் படுத்தப்பட்டு பின்னர் சொல்லடைவுகளாகவும் தொடரடைவுகளாகவும் பொருளடைவுகளாகவும் தொகுக்கப்பட்டன. ஆரம்ப கட்ட தமிழ் மொழியியல் காலகட்டத்தில் புறநானூறு, அகநானூறு, பதிற்றுப்பத்து, கலித்தொகை, கம்பராமாயணம், திருக்குறள் போன்ற நூல்களுக்கு வருணனை இலக்கணம் வரைவது நோக்கமாக அமைந்தது. அவ்வாறு செய்கையில் அந்நூல்களில் உள்ள சொற்கள் யாவும் மேற் சொன்ன முறையில் அகரவரிசை அட்டவணைகளாக தொகுக்கப்படுவது முதற்படி பணியாக அமைந்தது. பின்னர் அவற்றின் இலக்கணங்கள் ஆயப்பட்டு வருணனை மொழியியல் அடிப்படையில் விளக்கப்பட்டன. 1967-1970 காலகட்டங்களில் பிரஞ் இன்ஸ்டிடியூ ஓப் இண்டாலஜி, பாண்டிச்சேரி-இல் இருந்து வெளிவந்த சங்க இலக்கியங்களின் சொல்லடைவு (தற்போது 2007 பதிப்பாக A word Index of Cangam Literature என்ற பெயரில் Institute of Asian Studies-ஆல் வெளியிடப்பட்டுள்ளது.) என்பதும் மனித முயற்சியால் அட்டவணைப் படுத்தப்பட்டதுதான்.

சொல்லடைவு தயாரிப்பதையும் கணிப்பொறி மூலம் விரைவில் செய்துவிட இயலும். சொற்களையும் அவற்றோடு தொடர்பான செய்திகளையும் தந்தால் குறைந்த கால நேரத்தில் கணிப்பொறியின் உதவியால் சொல்லடைவு தயாரித்து விடலாம். இதன் படி நூல், முதலியவற்றில் ஒரு தனிச்சொல் பயின்று வரும் அனைத்து இடங்களையும் காட்டக்கூடியதாகக் கணிப்பொறியால் உருவாக்கப்பட்ட பட்டியல் கணினிவழி நூல் சொல்லடைவு எனப்படும்.

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

எடுத்துக்காட்டாகச் சங்க இலக்கியங்களுக்குச் சொல்லடைவு தயாரிப்பதை எடுத்துக் கொள்வோம். பின்வரும் அடிகள் சங்க இலக்கியமான புறநானூற்றிலிருந்து எடுக்கப்பட்டவை.

1. இவனியாரென்குவை யாயி னிவனே (புறம் 83.1)
2. தூங்குகையா னோங்க் நடைய (புறம் 22.1)

இவ்வடிகள் சொற்களாகப் பின்வருமாறு பிரித்துத் தரப்படவேண்டும். செயல்நிரல்கள் (programs) மூலம் இச்சொற்களை அகரவரிசைப்படுத்தி அவை ஒவ்வொன்றும் வரும் இடங்களைக் குறிப்பிட்டுத்தர இயலும்.

ஆயின் (புறம் 83.1)

இயார் (புறம் 83.1)

இவன் (புறம் 83.1)

என்குவை (புறம் 83.1)

ஏ (புறம் 83.1)

ஓங்கு (புறம் 22.1)

கையால் (புறம் 22.1)

தூங்கு (புறம் 22.1)

நடைய (புறம் 22.1)

ஒரு சொல் சங்க இலக்கியத்தில் எத்தனை முறை கையாளப் பெறுகின்றது அல்லது குறிப்பிட்ட ஒரு நூலில் எத்தனை முறை கையாளப் பெறுகின்றது என்பன போன்ற தகவல்களை செயல்நிரல்கள்/வழியமைப்புகள் மூலம் பெறலாம். தமிழ்ப்பல்கலைக்கழக அகராதியியல் துறையும் கணிப்பொறி அறிவியல் துறையும்

இணைந்து கணிப்பொறியைப் பயன்படுத்திச் சங்க இலக்கியங்களுக்கான ஒரு சொல்லடைவைத் தயாரித்து வெளியிட்டுள்ளன.

இராசேந்திரன் தமிழ்ப் பல்கலைக்கழகத்தால் வெளியிடப்பட்ட தமது 'தற்காலத் தமிழ்ச் சொற்களஞ்சியம்' என்ற நூலின் இறுதியில் சொற்களஞ்சியத்தில் வரும் சொற்களை அகரவரிசைப் படுத்தி அவை சொற்களஞ்சியத்தில் வரும் பக்கங்களை நோக்கீட்டுக் குறிப்புகளாகத் தரக் கணிணித் தொழில் நுட்பத்தைப் பயன்படுத்தினார்.

நூல் முதலியவற்றில் சொற்கள் பயின்று வரும் இடங்களையும் முறையையும் காட்டக்கூடிய விதத்தில் தொகுக்கப்பட்ட அச்சொற்களின் அகர வரிசைப் பட்டியல் சொற்பயில்வடங்கல் என்றும் சொல்லாட்சி அகராதி என்றும் அழைக்கப்படுகின்றது. சொல்லாட்சி அகராதி (Concordance) ஒரு பனுவலில்/உரையில் இடம் பெறும் சொற்களை அகரவரிசைப்படுத்தித் தருவதோடு அச்சொற்கள் நூலில் இடம்பெறும் இடங்கள், எத்தனை முறை இடம்பெறுகின்றன என்ற கணக்கீடு, எந்தெந்தச் சூழலில் பயன்படுத்தப்படுகின்றன என்பன போன்ற செய்திகளைத் தரும். சொற்பொருளடைவு (Compendium) சொல்லாட்சி அகராதிகளினின்றும் ஒருபடி மேற்சென்று சொற்கள் வரும் இடங்களையும் சூழல்களையும் பொருள் அடிப்படையில் பிரித்துத் தரும். கணிப்பொறியைப் பொறுத்தவரையில் இரண்டையும் ஒன்றில் அடக்கித் தந்துவிடலாம்.

புறநானூற்றிலிருந்து எடுக்கப்பட்ட ஒரு சொல்லின் சொற்பொருளடைவின் மாதிரி கீழே தரப்பட்டுள்ளது.

அணி

1. அழகு

1.1 அழகு தாழ்ந்த அணி (புறம் 57.1)

1.2 அணிகொள் தேர் (புறம் 68.14)

...

...

...

2. அணிகலன்

2.1. இழையணிபொலிந்த

(புற, 81.1)

...

...

...

3. ஒப்பனை

3.1. ஆடு கோட்பாட்டுச் சேரலாதன் போந்தைப் பொழியின் கண் அணியல்

(படி 51.9)

4. அணிமை

5. வரிசை

6. திரட்சி

7. ஆடை

8. ஒப்புமை

9. சூழ்தல்

10. பக்கரை (சேணம்)

சொல்லடைவு தயாரிக்க முன்னர் கூறியது போல் செய்யுளின் அடிகளைச் சொற்களாகப் பிரித்துத் தரவேண்டும். இத்துடன் சூழல் அடையாளப் படுத்தப்பட வேண்டும். செயல்நிரல்கள்/வழியமைப்புகள் மூலம் நாம் சொல்லாட்சி அகராதியையோ சொற்பொருடைவையோ தயாரிக்க இயலும். தமிழ்ப்பல்கலைக்கழக அகராதியியல் துறையும் கணிப்பொறி அறிவியல் துறையும் இணைந்து கணிப்பொறியைப்

பயன்படுத்திச் சங்ககால இலக்கியங்களுக்கு இத்தகைய சொல்லாட்சி அகராதியும் சொற்பொருளடைவும் தயாரித்துள்ளன.

2 அகராதியும் அகராதி உருவாக்கமும்

அகராதியியல்

அகராதியியல் என்பது அகராதி பற்றிய ஆய்வு, மேலும் இது இரண்டு தனித்தனிக் கல்வித் துறைகளாகப் பிரிக்கப்பட்டுள்ளது. இது அகராதிகளைத் தொகுக்கும் கலை. இது நடைமுறை அகராதியியல் (Practical lexicography), கோட்பாட்டு அகராதியியல் (Theoretical lexicography) என இருவகைப்படும். நடைமுறை அகராதியியல் என்பது அகராதிகளைத் தொகுத்தல், எழுதுதல் மற்றும் திருத்துதல் ஆகியவற்றின் கலை அல்லது கைவினையாகும். கோட்பாட்டு அகராதியியல் என்பது ஒரு மொழியின் அகராதியின் (சொற்றொகை) சொல்லன்களின் (lexemes) எழுத்துவடிவம் (orthographic) பொருண்மையியல் (semantic), உறுப்பமைவு (syntagmatic) மற்றும் அடுக்குநிலை (paradigmatic) பண்புக்கூறுகள் பற்றிய அறிவார்ந்த ஆய்வு ஆகும், மேலும் அகராதிகளில் உள்ள தரவை இணைக்கும் அகராதிக் கூறுகள் மற்றும் கட்டமைப்புகளின் கோட்பாடுகளை உருவாக்குதல், குறிப்பிட்ட வகைகளில் பயனர்களின் தகவலுக்கான தேவைகள். சூழ்நிலைகள், மற்றும் அச்சிடப்பட்ட மற்றும் மின்னணு அகராதிகளில் உள்ள தரவைப் பயனர்கள் எவ்வாறு சிறப்பாக அணுகலாம் என்பது பற்றியாகும். இது சில நேரங்களில் 'கருவி அகராதியியல்' (metalexicography) என்று குறிப்பிடப்படுகிறது.

அகராதியியலில் இருந்து வேறுபட்ட, சொல்லியியலின் (lexicology) வரையறையில் சில கருத்து வேறுபாடுகள் உள்ளன. சிலர் "சொல்லியல்" என்பது கோட்பாட்டு அகராதியியலுக்கு ஒத்த பொருளாகப் பயன்படுத்துகின்றனர்; மற்றவர்கள் அதை ஒரு குறிப்பிட்ட மொழியில் உள்ள சொற்களின் பட்டியல் தொடர்பான மொழியியலின் ஒரு பிரிவாகப் பயன்படுத்துகின்றனர். அகராதியியலில் ஈடுபாடு கொண்ட ஒருவர், அகராதியியலார் என்று அழைக்கப்படுகிறார்.

பொது அகராதிகளின் வடிவமைப்பு, தொகுத்தல், பயன்பாடு மற்றும் மதிப்பீடு ஆகியவற்றின் மீது பொது அகராதியியல் கவனம் செலுத்துகிறது, அதாவது பொதுவான பயன்பாட்டில் உள்ள மொழியின் விளக்கத்தை வழங்கும் அகராதிகள். இத்தகைய அகராதி பொதுவாக பொது அகராதி அல்லது பொது நோக்கத்திற்கான மொழி (Language for General Purpose (LGP)) அகராதி என்று அழைக்கப்படுகிறது. சிறப்பு அகராதிகளின் [அதாவது ஒன்று அல்லது அதற்கு மேற்பட்ட சிறப்புப் பாடத் துறைகளின் (ஒப்பீட்டளவில் கட்டுப்படுத்தப்பட்ட) மொழியியல் மற்றும் உண்மைக் கூறுகளின் தொகுப்பிற்கு அர்ப்பணிக்கப்பட்ட அகராதிகள், எ.கா. சட்ட அகராதி] வடிவமைப்பு, தொகுத்தல், பயன்பாடு மற்றும் மதிப்பீடு ஆகியவற்றில் கவனம் செலுத்துகிறது. அத்தகைய அகராதி பொதுவாகச் சிறப்பு அகராதி (specialized dictionary) அல்லது குறிப்பிட்ட நோக்கங்களுக்கான மொழி அகராதி (Language for specific purposes dictionary) என அழைக்கப்படுகிறது மற்றும் நீல்சன் 1994-ஐப் (Nielsen 1994), பின்பற்றி, சிறப்பு அகராதிகள் பல்புல (multi-field), ஒற்றைப்புல (single-field) அல்லது துணைப்புல (sub-field) அகராதிகளாகும்.

அகராதியியலில் முதன்மையான ஆய்வுப் பொருள் அகராதியாக இருப்பதால், அகராதியியலானது அதன் சொந்த உரிமையில் ஒரு புலமைசார் துறையே தவிர, பயன்பாட்டு மொழியியலின் (applied linguistics) துணைப் பிரிவு அல்ல என்பது இப்போது பரவலாக ஏற்றுக்கொள்ளப்பட்டுள்ளது (எ.கா. பெர்கன்ஹோல்ட்ஸ்/நீல்சன்/டார்ப் Bergeholtz/Nielsen/Tarp 2009ஐப் பார்க்கவும்).

அகராதியிலார்கள் எளிமையான சொற்களை வரையறுப்பதோடு, கூட்டு அல்லது சிக்கலான சொற்கள் அல்லது பல அர்த்தங்களைக் கொண்ட சொற்கள் எவ்வாறு தெளிவாக விளக்கப்படலாம் என்பதைக் கண்டறிவதில் பணிபுரிகின்றனர். அகராதியில் எந்தச் சொற்களை வைக்க வேண்டும், சேர்க்க வேண்டும் அல்லது நீக்க

வேண்டும் என்பது குறித்தும் அவர்கள் முடிவு செய்கிறார்கள். புரிந்துகொள்வதற்கும் வழிச்செலுத்துவதற்கும் வசதியாக சொல்சார் பொருட்களை (பொதுவாக அகரவரிசைப்படி) ஒழுங்கமைப்பது அவர்கள் பொறுப்பு.

நடைமுறை அகராதி வேலை பல செயல்பாடுகளை உள்ளடக்கியது, மேலும் நன்கு வடிவமைக்கப்பட்ட அகராதிகளின் தொகுப்பிற்கு பின்வரும் அனைத்து அல்லது சில அம்சங்களையும் கவனமாகப் பரிசீலிக்க வேண்டும்:

- நோக்கம் கொண்ட பயனர்களின் விவரக்குறிப்பு (அதாவது மொழியியல் மற்றும் மொழியியல் அல்லாத திறன்கள்) மற்றும் அவர்களின் தேவைகளை அடையாளம் காணுதல்
- அகராதியின் கருத்துப்பரிமாற்றம் மற்றும் புலனறிவு செயல்பாடுகளை வரையறுத்தல்
- அகராதியின் கூறுகளைத் தேர்ந்தெடுத்து ஒழுங்கமைத்தல்
- அகராதியில் தரவை வழங்குவதற்கு பொருத்தமான கட்டமைப்புகளைத் தேர்ந்தெடுப்பது (அதாவது சட்ட அமைப்பு/frame structure, விநியோக அமைப்பு/distribution structure, பருமக் கட்டமைப்பு/macro-structure, நுண் கட்டமைப்பு/micro-structure மற்றும் குறுக்கு-குறிப்பு அமைப்பு/cross-reference structure)
- உள்ளீடுகளாக முறைப்படுத்தலுக்கான சொற்கள் மற்றும் இணைப்புகளைத் தேர்ந்தெடுப்பது
- சொல்லடை வகைப்பாடுகள் (collocations), சொற்றொடர்கள் மற்றும் எடுத்துக்காட்டுகள் தேர்வு
- ஒவ்வொரு சொல்லுக்கும் அல்லது சொல்லின் ஒரு பகுதிக்கும் சொல்லன் வடிவங்களைத் (lemma forms) தேர்வு செய்தல்
- சொரற்களை வரையறுத்தல்

- வரையறைகளை ஒழுங்கமைத்தல்
 - சொற்களின் உச்சரிப்புகளைக் குறிப்பிடுதல்
 - பொருத்தமான இடங்களில் கட்டுண்டவழக்கு/register மற்றும் பேச்சுவழக்குக்கான/dialect வரையறைகள் மற்றும் உச்சரிப்புகளை அடையாளப்படுத்தல்,
 - இரு மொழி மற்றும் பல மொழி அகராதிகளில் நிகரன்களைத் தேர்ந்தெடுப்பது
 - இரு மொழி மற்றும் பன்மொழி அகராதிகளில் சொல்லடிவகைப்பாடு (collocations), சொற்றொடர்கள் மற்றும் எடுத்துக்காட்டுகளை மொழிபெயர்த்தல்
 - அச்சிடப்பட்ட மற்றும் மின்னணு அகராதிகளில் உள்ள தரவைப் பயனர்கள் அணுகுவதற்கான சிறந்த வழியை வடிவமைத்தல்
- அகராதி பயனர்களால் ஏற்படும் அகராதி தகவல் செலவுகளை முடிந்தவரை குறைவாக வைத்திருப்பது அகராதியின் ஒரு முக்கிய குறிக்கோள். நீல்சன் (Nielsen 2008) அகராதிகளை உருவாக்கும் போது அகராதியாளர்கள் கருத்தில் கொள்ள வேண்டிய பொருத்தமான அம்சங்களைப் பரிந்துரைக்கிறார், ஏனெனில் அவை அனைத்தும் பயனர்களின் தோற்றத்தையும் குறிப்பிட்ட அகராதிகளின் உண்மையான பயன்பாட்டையும் பாதிக்கின்றன.

கோட்பாட்டு அகராதியியல் என்பது அகராதியியலின் அதே அம்சங்களைப் பற்றியது, ஆனால் எதிர்கால அகராதிகளின் தரத்தை மேம்படுத்தக்கூடிய கொள்கைகளை உருவாக்குவதை நோக்கமாகக் கொண்டுள்ளது, எடுத்துக்காட்டாக, தரவு அணுகல் மற்றும் அகராதி தகவல் செலவுகள். இத்தகைய கல்வி அகராதி ஆராய்ச்சியின் பல முன்னோக்குகள் அல்லது கிளைகள் வேறுபடுகின்றன: 'அகராதி விமர்சனம்' (அல்லது ஒன்று அல்லது அதற்கு மேற்பட்ட அகராதிகளின் தரத்தை மதிப்பீடு செய்தல், எ.கா. மதிப்புரைகள் மூலம் (பார்க்க நீல்சன்/Nielsen 1999), 'அகராதி வரலாறு' (அல்லது

ஒரு குறிப்பிட்ட நாட்டில் அல்லது மொழியில் உள்ள அகராதியின் வகை அல்லது ஒரு மரபுகளைக் கண்டறிதல்), 'அகராதி வகைப்பாடியியல்' (அல்லது அகராதி மற்றும் கலைக்களஞ்சியம், ஒருமொழி மற்றும் இருமொழி அகராதி, பொது மற்றும் தொழில்நுட்ப அல்லது கல்வியியல் அகராதி போன்ற பல்வேறு வகையான நோக்கீட்டுப் படைப்புகளை வகைப்படுத்துதல்), 'அகராதி அமைப்பு ' (அல்லது அகராதியில் தகவல்களை வழங்குவதற்கான பல்வேறு வழிகளை வடிவமைத்தல்), 'அகராதி பயன்பாடு' (அல்லது அகராதி பயனர்களின் குறிப்புச் செயல்கள் மற்றும் திறன்களைக் கவனித்தல்), மற்றும் 'அகராதி தகவல் தொழில்நுட்பம்' (அல்லது அகராதியின் தொகுப்புச் செயல்முறைக்குக் கணினி உதவிகளைப் பயன்படுத்துதல்).

ஒரு முக்கியமான கருதல் 'இருமொழி அகராதியின்' நிலை, அல்லது இருமொழி அகராதியை அதன் அனைத்து அம்சங்களிலும் தொகுத்தல் மற்றும் பயன்படுத்துதல் (எ.கா. நீல்சன்/Nielsen 1894-ஐப் பார்க்கவும்). இந்த வகை அகராதிக்கு ஒப்பீட்டளவில் நீண்ட வரலாறு இருந்தபோதிலும், அதன் ஒருமொழிய இணையைக் காட்டிலும் பல விஷயங்களில் இது குறைவாகவே வளர்ந்ததாகக் கூறப்படுகிறது, குறிப்பாக, சம்பந்தப்பட்ட மொழிகளில் ஒன்று முக்கிய மொழியாக இல்லாத சந்தர்ப்பங்களில். அனைத்து வகையான குறிப்புப் படைப்புகளும் மொழிகளுக்கிடையேயான பதிப்புகளில் கிடைக்காது, எ.கா. எல்எஸ்பி, கற்றவர்கள் மற்றும் கலைக்களஞ்சிய வகைகள், சில நேரங்களில் இந்த சவால்கள் புதிய துணை வகைகளை உருவாக்குகின்றன, எ.கா. Hornby's (Oxford) Advanced Learner's Dictionary English-Chinese போன்ற 'அரை-இருமொழி' அல்லது 'இருமொழி' அகராதிகள், அவை ஏற்கனவே உள்ள ஒருமொழி அகராதிகளை மொழிபெயர்த்து உருவாக்கப்பட்டுள்ளன (பார்க்க Marelllo 1998).

கணிப்பொறி அகராதி உருவாக்குதல்

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

கணினிச் சொல்லியல் (computational lexicology) என்பது அகராதியைப் பற்றி படிப்பதற்குக் கணினியைப் பயன்படுத்துவதைப் பற்றிக் கூறுகின்ற கணினிமொழியியலின் கிளை ஆகும். இயந்திரம் படிக்கவியலும் அகராதிகளைப் பற்றி படிப்பதுதான் கணினி சொல்லியல் எனக் குறுகலாகவும் வரையறை விளக்கம் செய்வர். இது கணினி அகராதியல் என்பதிலிருந்து வேற்படுத்தப்படவேண்டும். கணினி அகராதியியல் என்பது கணியைப் பயன்படுத்தி அகராதியை உருவாக்குவதை குறிப்பிடும்.

அகராதியியலும் மொழியியலும்

மொழியியல் அகராதியியலாரின் அகராதியைச் செம்மைப்படுத்தப் பின்வரும் வழிகளின் பயன்படுகிறது:

- பொருண்மை வேறுபாடுகளின் ஒழுங்கமைப்பைக் கண்டுகொள்ள
- தொடரியல் பண்புகளையும் பொருண்மையியல் பண்புகளையும் பொருத்த
- ஒரே கருத்துரு வகுப்பைச் சார்ந்த சொல் பொருண்மைகளை விவரணை செய்யப் பொதுவான பொருண்மைச் சட்டத்தைக் கையாள

ஒரு அகராதி (dictionary மற்றும் lexicon என்ற சொற்கள் ஒருபொருள் பன்மொழிகளாகப் பயன்படுத்தப்படுவதுண்டு) என்பது ஒரு மொழியின் சொற்றொகை பற்றிய தகவல்களை வழங்கும் ஒரு சொல்சார் தரவுத் தளமாகும். இது மொழியின் ஒவ்வொரு சொல்லுக்கும் அதன் சொல்வகைப்பாடு, இலக்கண வகைப்பாடு மற்றும் தொடரியல்-பொருண்மையியல் அம்சங்கள் மற்றும் சூழல் கட்டுண்ட தகவல் (context sensitive information) (பயன்வழியியல்சார்/pragmatic) ஆகியவற்றைக் குறிப்பிடுகிறது. dictionary என்ற சொல் லத்தீன் வார்த்தையான dictionarium என்பதிலிருந்தும் மற்றும்

lexicon என்ற வார்த்தை கிரேக்கச் சொல்லான lexicon என்பதிருந்தும் பெறப்பட்டன. அகராதி தகவல்களை நேரடியாக வழங்கும் பல வளங்கள் உள்ளன. இவை சொற்களஞ்சியம், தொடரடைவு, சொற்களஞ்சியம். கலைக்களஞ்சியம் முதலியன.

கணினிசார் அகராதியியல் (computational lexicography) என்பதும் கணினிசார் சொல்லியியல் (computational lexicology) என்பதும் கணினி மொழியியலின் (Computational linguistics) ஒரு கிளை ஆகும்; இவை அகராதி ஆய்வில் கணினிகளைப் பயன்படுத்துவதில் அக்கறை கொண்டுள்ளது. இயந்திரம் படிக்கக்கூடிய அகராதிகளின் ஆய்வில் கணினிகளின் பயன்பாடு என சில அறிஞர்கள் (Amsler, 1980) இதை மிகவும் சுருக்கமாக விவரித்திருக்கிறார்கள். கணினிசார் சொல்லியியல் கணினிசார் அகராதியியலிருந்து (computational lexicography) வேறுபடுகிறது; அகராதிகளைத் நிர்மாணிப்பதில் கணினிகளைப் பயன்படுத்துவது எனக்கூறுவது மிகவும் சரியாக இருக்கும்; இருப்பினும் சில ஆராய்ச்சியாளர்கள் கணினிசார் அகராதியியலைக் கணினிசார் சொல்லியலின் ஒருபொருள்பன்மொழியாகப் பயன்படுத்துகின்றனர்

வரலாறு

1960களில் கணினி ஒழுங்குமுறை வளர்ச்சி நிறுவனத்தில் ஜான் ஒல்னி மற்றும் பலரால் மெரியம்-வெப்ஸ்டர் ஏழாவது கல்லூரி அகராதி (Merriam-Webster Seventh Collegiate Dictionary) மற்றும் மெரியம்-வெப்ஸ்டர் புதிய பாக்கெட் அகராதி (Merriam-Webster New Pocket Dictionary) ஆகியவற்றின் இயந்திரம் படிக்கக்கூடிய நாடாக்களை உருவாக்கத் தொடங்கி, இயந்திரத்தால் படிக்கக்கூடிய அகராதிகளின் தோற்றத்துடன் கணினி மொழியியலில் ஒரு தனி ஒழுக்கமாகக் கணினிசார் சொல்லியியல் வெளிப்பட்டது. இன்று, சொல்வலையின் (வேர்ட்நெட்டின்) (WordNet)

உருவாக்கம் மற்றும் பயன்பாடுகள் மூலம் கணினிசார் அகராதியியல் அறியப்படுகிறது. காலப்போக்கில் ஆராய்ச்சியாளர்களின் கணினியல்சார் செயலாக்கம் அதிகரித்ததால், உரை பகுப்பாய்வில் கணினிசார் சொல்லியலின் பயன்பாடு எங்கும் காணப்படுகிறது. 1987ஆம் ஆண்டில், பிறருடன் பைர்ட் (Byrd), கால்சோலாரி (Calzolari), சோடோரோ (Chodorow) என்போர் உரை பகுப்பாய்விற்கான கணினிசார் கருவிகளை (computational tools) உருவாக்கியுள்ளனர். குறிப்பாக பல்பொருளொருமொழிச் சொற்களின் (polysemous words) அர்த்தங்களை உள்ளடக்கிய சங்கங்களை ஒருங்கிணைக்க இந்த மாதிரி வடிவமைக்கப்பட்டுள்ளது.

சொல்லியியல்

அகராதி ஆய்வில் கவனம் செலுத்துகிற மொழியியலின் ஒரு பகுதிதான் சொல்லியல் (Lexicology). இது அகராதியின் இயல்பு, அமைப்பு மற்றும் அதன் அலகுகளின் குணங்களை உருப்படுத்தும் செய்யும்; இயற்கை மொழிகளின் மாதிரிகளில் அகராதியின் பங்கை விளக்கும். கணினி சொல்லியல் (computational lexical semantics) சொல் ஆய்வில் கணினியின் பயன்பாட்டின் நோக்கை உள்ளடக்குகின்றது. குறிப்பாகப் பின்வருவனவற்றில் கவனக்குவிப்பு செய்கின்றது:

- அகராதியின் கணினி உருப்படுத்தத்தில் சொல் தரவில் செய்யப்படும் கணினிச் செயல்பாடுகள்
- ஒரு நோக்கில் கணினி அகராதிகளுக்கிடையே உள்ள உறவுகள், மற்றொரு நோக்கில் பல இயற்கை மொழி ஆய்வு (NLP) ஒழுங்குமுறையின் கூறுகளுக்கிடையே உள்ள உறவுகள்

அகராதியியல்

அகராதியியல் சொற்களைச் சேகரிப்பது மற்றும் அவை எவ்வாறு பயன்படுத்தப்படுகின்றது என்பதில் கவனம் செலுத்துவதாகப் பொதுவாகக் கருதப்படுகின்றது. அகராதியியல் உற்பத்திப்பொருள்களில் பல அச்சடிக்கப்பட்ட அகராதிகள் மட்டுமன்றி அகரவரிசை அட்டவணை, சொல்லடைவு, தொடரடைவு, கலைச்சொற்கள் போன்றவையும் அடங்கும். அகராதியியல் பின்வரும் செயல்பாடுகளை உட்படுத்தும்:

- குறிப்பிடத்தகுந்த சொற்களையும் கலைச்சொற்களையும் தொடர்பான செய்திகளுடன் கண்டுபிடிப்பது,
- அம்மாதிரியான செய்திகளைத் தொகுப்பது, குறியாக்கம் செய்வது, முன்னிலைப்படுத்தித் தருவது,
- அம்மாதிரியான செய்திகளைப் பரவச்செய்தல் மற்றும் வழங்குதல், அந்தச் செய்திகளின் பயன்பாட்டை ஆய்தல் மற்றும் அதன் தரத்தை அளத்தல்.

அகராதி உருவாக்கல்

அகராதி உருவாக்கல் என்பது பன்னோக்கு செயற்பாங்காகும். இது குறைந்தது பின்வரும் செயற்பாடுகளை உள்ளடக்கும்:

- மொழியை/மொழிகளைப் புரிந்து கொள்ளுதல் (Understanding the language(s); அமைப்பு, செயற்பாடு, பொருண்மை மற்றும் சமுதாய-பண்பாட்டு அடிப்படையில் மொழியின் அமைப்பைப் புரிந்து கொள்ளுதல்.
- செய்திகளை அமைப்பாக்கம் செய்தல் (Structuring the information): அகராதிச் சொல்லின் கீழ் மேற்சொன்ன செய்திகளை ஒரு அமைப்பொழுங்கில் தருதல்.

- செய்திகளை உள்ளீடு செய்தல் (தரவு மையத்தைத் தொகுத்தல்) (Inputting the information (compiling the lexical database): இயல்பாக ஒரு ஆண்டிற்குச் செய்திகளை உள்ளீடு செய்தல், இது ஆரம்பத்திலேயே செய்யப்பட வேண்டும்; தொடர்ந்து செய்யப்பட வேண்டும்.
- பரிசோதித்து மேன்மைப்படுத்துதல் (Checking and refining): சொல் தரவு மையத்திலுள்ள செய்திகளைப் பரிசோதனை செய்து மேன்மைப் படுத்தல்.
- தரவைப் பயன்படுத்தல் (Manipulation of the data): தரவைச் சொற்களங்களைப் பெறுவதற்காகவும், தலைகீழான செயல்பாடுகளைச் செய்வதற்காகவும் பயன்படுத்தல்.
- வெளியீடு: வடிவமைப்பைத் தீர்மானித்து வேண்டிய மாற்றங்களைச் செய்தல்.
- அச்சிடல்.
- வணிகம் செய்தல் மற்றும் விநியோகித்தல்.

அகராதி வகைகள்

அகராதிகள் அவற்றின் அமைப்பு பயன்பாடு போன்றவற்றின் அடிப்படையில்

பின்வருமாறு வகைப்படுத்தப்படும்:

- ஒரு மொழி, இருமொழி, பன்மொழி அகராதிகள்
- பொருட்புல அகராதிகள் மற்றும் கருத்துரு அடிப்படையிலான அகராதிகள்
- தனிப்பட்ட வகை (கொச்சை மொழி, சொல்மூலம் அல்லது சொல் வரலாறு)
- கலைக்களஞ்சியம் (உண்மை உலகத்தின் குறிப்புரை)
- மொழிபெயர்ப்பு (மொழிபெயர்ப்பு நிகரிகள்)
- பேருந்தரவுகள் (தொடரடைவு)

- சொற்றொகை (சொற்களம் அடிப்படையிலான சொற் பட்டியல்)
- பொருள்விளக்கச் சொற்கோவை (பழைய, அரிதான சொற்கள்)
- பொருள்களுக்குப் பெயரிடுதல்
- சொல்லடைவு
- சொல், கலைச்சொல் நிகழ்வெண்

அகராதி ஆய்வு

கணினிசார் நோக்கங்களுக்காக அச்சு அகராதிகளின் உள்ளடக்கம் மற்றும் வரம்புகளைப் புரிந்துகொள்வதற்கு கணினிசார் அகராதி பங்களிப்பு செய்துள்ளது (அதாவது, கணினிசார் மொழியியலின் தேவைகளுக்கு முந்தைய சொற்களஞ்சியம் போதுமானதாக இல்லை என்பதை இது தெளிவுபடுத்தியது). கணினிசார் சொல்லியல் ஆய்வாளர்களின் பணியின் மூலம் அச்சு அகராதிப் பதிவின் ஒவ்வொரு பகுதியும் பின்வருவனவற்றிலிருந்து ஆய்வு செய்யப்பட்டுள்ளன:

1. எது ஒரு தலைப்புச்சொலை அமைக்கிறது - எழுத்துத் திருத்தப் பட்டியல்களை உருவாக்கப் பயன்படுகிறது;
2. தலைப்புச்சொல் என்ன மாறுபாடுகள் மற்றும் திரிபுகளை உருவாக்குகிறது - உருபனியல் அமைப்பை அனுபவ ரீதியாகப் புரிந்துகொள்ளப் பயன்படுகின்றது;
3. தலைப்புச்சொல் எவ்வாறு அசைகளாக பிரிக்கப்பட்டுள்ளது;
4. தலைப்புச்சொல் எவ்வாறு உச்சரிக்கப்படுகிறது - பேச்சு உருவாக்க ஒழுங்குமுறைகளில் பயன்படுத்தப்படுகிறது;

5. தலைப்புச்சொல் எடுக்கும் சொல்வகைப்பாடு - சொல்வகைப்பாட்டு அடையாளப்படுத்திகளுக்குப் (parts of speech/POS taggers) பயன்படுத்தப்படுகிறது;
6. தலைப்புக்கு ஒதுக்கப்பட்ட எந்தச் சிறப்புப் பொருள் அல்லது பயன்பாட்டுக் குறியீடுகள் - உரை ஆவணப் பொருள் விஷயங்களை அடையாளம் காணப் பயன்படுகிறது;
7. தலைப்புச்சொல்லின் வரையறைகள் மற்றும் அவற்றின் தொடரியல் - சூழலில் சொல்லின் பொருண்மை மயக்கம் நீக்க ஒரு உதவியாகப் பயன்படுத்தப்படுகிறது;
8. தலைப்புச்சொல்லின் சொற்பிறப்பியல் மற்றும் சொற்றொகையைத் தோற்றுவிக்கும் மொழிகளால் பண்பக்கம்செய்ய அதன் பயன்பாடு - உரைச் சொற்றொகையை அதன் தோற்ற மொழிகளாக வகைப்படுத்தப் பயன்படுகிறது;
9. எடுத்துக்காட்டு வாக்கியங்கள்;
10. தலைப்பிலிருந்து உருவாகும் கூடுதல் சொற்கள் மற்றும் பல் சொல் வெளிப்பாடுகள்; மற்றும்
11. ஒருபொருள் பன்மொழிகள் மற்றும் எதிர்மொழிகள் போன்ற தொடர்புள்ள சொற்கள்.

கணினி நிரல்களுக்கான போதுமான தொடரியல் மற்றும் பொருண்மையியல் தகவல்கள் இல்லாததால் கணினி மொழியியலுக்கான மூலவளமாக அச்சு அகராதிகளைக் கருத இயலாமல், பல கணினி மொழியியலாளர்கள்

ஏமாற்றமடைந்தனர். கணினிசார் சொல்லியல் பற்றிய பணி விரைவாக இரண்டு கூடுதல் திசைகளில் முயற்சிகளுக்கு வழிவகுத்தது.

கணினிசார் சொல்லியலுக்கு வாரிசுகள்

முதலாவதாக, கணினி மொழியியலாளர்களுக்கும் அகராதியியலார்களுக்கும் இடையிலான கூட்டு நடவடிக்கைகள் அகராதிகளை உருவாக்குவதில் தரவுத்தொகுதி வகித்த பங்கைப் புரிந்துகொள்ள வழிவகுத்தது. பெரும்பாலான கணினிசார் சொல்லியலார்கள் அகராதிகளை உருவாக்க அகராதி வல்லுநர்கள் பயன்படுத்திய அடிப்படைத் தரவுகளைச் சேகரிக்க பெரிய தரவுத்தொகுதிகளை உருவாக்க முயன்றனர். தரவு சேகரிப்பு முயற்சி (Data Collection Initiative (ACL/DCI ஏ.சி.எல்/டி.சி.ஐ) மற்றும் மொழியியல் தரவுக் கூட்டமைப்பு (Linguistic Data Consortium (LDC/எல்.டி.சி)) ஆகியவை இந்தப் பாதையில் சென்றன. மார்க்அப் மொழிகளின் வருகை அடையாளப்படுத்தப்பட்ட தரவுத்தொகுதிகளை உருவாக்க வழிவகுத்தது; இது கணினி மொழியியல் ஒழுங்குமுறைகளை உருவாக்க எளிதாகப் பகுப்பாய்வு செய்ய முடியும். சொல்வகைப்பாட்டு அடையாளப்படுத்தப்பட்ட தரவுத்தொகுதிகள் மற்றும் பொருண்மை அடையாளப்படுத்தப்பட்ட தரவுத்தொகுதிகள் என்பன சொல்வகை அடையாளப்படுத்திகள் மற்றும் சொல்மயக்கநீக்கத் தொழில்நுட்பம் இவற்றைச் சோதித்து மேம்படுத்துவதற்காக உருவாக்கப்பட்டன.

இரண்டாவது திசை சொல்சார் அறிவுத் தளங்களை (Lexical Knowledge Bases (LKBS /எல்.கே.பி) உருவாக்குவதை நோக்கி இருந்தது. கணினி மொழியியல் நோக்கங்களுக்காக, குறிப்பாக கணினி சொற்பொருண்மையியல் நோக்கங்களுக்காக ஒரு அகராதி என்னவாக இருக்க வேண்டும் என்று ஒரு சொல்சார் அறிவுத் தளம்

கருதப்பட்டது. இது ஒரு அச்ச அகராதியில் உள்ள அதே தகவலைக் கொண்டிருக்க வேண்டும்; ஆனால் சொற்களின் பொருண்மைகள் மற்றும் அர்த்தங்களுக்கு இடையிலான பொருத்தமான இணைப்புகள் குறித்து முற்றிலும் விளக்கப்பட்டுள்ளது. கணினி பகுப்பாய்வில் பயன்படுத்த அகராதிகள் உருவாக்கப்பட்டிருந்தால், பலர் விரும்பிய வளங்களை உருவாக்கத் தொடங்கினர். ஃபில்மோர் இன் ஃபிரேம்நெட் வேலை (FrameNet work) போன்ற தொடரியல் மற்றும் பொருண்மைத் தகவல்களை விவரிக்கும் புதிய முயற்சிகள் போலவே, சொல்வலை (வேர்ட்நெட்/WordNet) அத்தகைய வளர்ச்சியாக கருதப்படுகிறது. கணினி மொழியியலுக்கு வெளியே, செயற்கை நுண்ணறிவின் ஒன்டாலஜி பணி செயற்கை அறிவுப் பயன்பாடுகளுக்கான ஒரு சொல்சார் அறிவுத் தளத்தை உருவாக்குவதற்கான ஒரு பரிணாம முயற்சியாகக் காணலாம்.

தரப்படுத்தல்

கணினி அகராதிகளின் உருவாக்கம், பராமரிப்பு மற்றும் நீட்டிப்பை மேம்படுத்துவது இயற்கைமொழி ஆய்வை (NLP.) பாதிக்கும் முக்கியமான அம்சங்களில் ஒன்றாகும். முக்கிய சிக்கல் இயங்குதன்மை: பல்வேறு அகராதிகள் அடிக்கடி பொருந்தாது. மிகவும் அடிக்கடி வரும் நிலைமை: இரண்டு அகராதிகள் அல்லது அகராதிகளின் துண்டுகளை எவ்வாறு இணைப்பது? இரண்டாம் நிலை சிக்கல் என்னவென்றால், ஒரு அகராதி பொதுவாக ஒரு குறிப்பிட்ட இயற்கை மொழி ஆய்வு (NLP/என்.எல்.பி.) திட்டத்திற்கு ஏற்ப வடிவமைக்கப்பட்டுள்ளது மற்றும் பிற இயற்கை மொழி ஆய்வுத் திட்டங்கள் அல்லது பயன்பாடுகளுக்குள் பயன்படுத்துவதில் சிரமங்கள் உள்ளன.

அகராதி உருவாக்குவதற்கான தேவைகள்

அகராதிகள் ஏன் முக்கியம்?

உங்களுக்கு புரியாத எந்த வார்த்தைகளின் அர்த்தத்தையும் அறிய நீங்கள் ஒரு அகராதியைப் பயன்படுத்தலாம். இருப்பினும், பல்வேறு வகையான அகராதிகள் நிறைய உள்ளன, எனவே நீங்கள் எதைப் பயன்படுத்த வேண்டும்? என்ற கேள்வி எழும்.

உங்கள் படிப்பின் போது உங்களுக்கு மிகவும் உதவியாக இருக்கும் சில வகையான அகராதிகளைப் பார்ப்போம்.

•ஒரு பல்கலைக்கழகத்தில் நீங்கள் படிக்கும் நேரத்தில் மிக முக்கியமான கருவிகளில் ஒன்று அகராதி. ஒரு நல்ல அகராதி உங்கள் விஷயத்தை நன்கு புரிந்துகொள்ளவும், உங்கள் தகவல்தொடர்புகளை மேம்படுத்தவும், நீங்கள் தரங்களை சரியாகப் பயன்படுத்துகிறீர்கள் என்பதை உறுதிப்படுத்துவதன் மூலம் உங்கள் தரங்களை மேம்படுத்தவும் உதவும்.

•இருப்பினும், எல்லா அகராதிகளும் ஒரே மாதிரியானவை அல்ல, அகராதியை சரியாகப் பயன்படுத்துவது உங்களுக்குத் தெரியாவிட்டால், அது உண்மையில் ஒரு வார்த்தையின் தவறான பொருளைக் கற்பிக்கும் மற்றும் நல்ல தரத்தைப் பெறுவது மிகவும் கடினம்.

•கீழே உள்ள நான்கு வெவ்வேறு வகையான அகராதிகளைப் பாருங்கள். புதிய சொல்லை நீங்கள் புரிந்து கொள்ளாதபோது எந்த அகராதி சிறந்த தேர்வு என்று நினைக்கிறீர்கள்?

இருமொழி அகராதி

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

இருமொழி அகராதி என்பது உங்கள் சொந்த மொழியில் மொழிபெயர்க்க நீங்கள் தேடும் வார்த்தையைக் கொண்ட ஒரு அகராதி. இருமொழி அகராதி பொதுவாக நீங்கள் தேடும் சொற்களைப் பற்றிய சிறிய அளவிலான தகவல்களை மட்டுமே கொண்டிருக்கிறது, ஆனால் அதை புரிந்துகொள்வது எளிதானது, ஏனெனில் இது உங்கள் சொந்த மொழியில் எழுதப்பட்டுள்ளது.

ஒருமொழி அகராதி

ஒரு மொழியியல் அகராதி ஒருமொழியில் மட்டுமே எழுதப்பட்டுள்ளது. ஒரு ஒருமொழி அகராதியில் மூலமொழியில் உள்ள ஒவ்வொரு சொல்லையும் பற்றிய பல்வேறு தகவல்கள் உள்ளன. இந்த அகராதிகளில் இலக்கணம் மற்றும் உச்சரிப்பு பற்றிய ஏராளமான தகவல்கள் உள்ளன. ஒரு ஒருமொழி அகராதி சில நேரங்களில் குழப்பமானதாக இருக்கலாம், ஏனெனில் அது ஒவ்வொரு சொல்லுக்கும் பல வரையறைகளைக் கொண்டுள்ளது.

கற்றல் அகராதி

ஒரு கற்றல் அகராதி ஒரு மொழி அகராதியைப் போன்றது, ஏனெனில் இது இயல்மொழியில் மட்டுமே உள்ளது மற்றும் நீங்கள் கண்டுபிடிக்க முயற்சிக்கும் சொல்லைப் பற்றி நிறைய தகவல்கள் உள்ளன. இருப்பினும், ஒரு கற்றல் அகராதி ஒரு நிலையான ஒருமொழி அகராதியைக் காட்டிலும் பயன்படுத்த எளிதானது மற்றும் தெளிவான எடுத்துக்காட்டுகள் மற்றும் எளிமைப்படுத்தப்பட்ட மொழியைக் கொண்டுள்ளது.

ஒரு பொருள் சார்ந்த அகராதி

ஒரு பொருள் சார்ந்த அகராதி ஒரு ஒருமொழி அகராதியைப் போன்றது, ஆனால் அதில் சொற்களின் பொதுவான அர்த்தம் இல்லை, குறிப்பிட்ட பாடத்திற்கு பொருத்தமான வரையறைகள் மட்டுமே. உங்கள் பொருள் பகுதியில் ஒரு சொல் எவ்வாறு வித்தியாசமாக பயன்படுத்தப்படுகிறது என்பதை நீங்கள் கண்டுபிடிக்க வேண்டியிருக்கும் போது பொருள் சார்ந்த அகராதிகள் மிகவும் பயனுள்ளதாக இருக்கும், ஆனால் அவை பயன்படுத்த மிகவும் கடினமாக இருக்கும், மேலும் நீங்கள் தேடும் சொல்லைக் கொண்டிருக்கக்கூடாது.

அகராதியைப் பயன்படுத்துவதற்கான காரணங்கள்:

புதிய மொழியைக் கற்கும் எவருக்கும் அகராதி மிக முக்கியமான கருவியாகும்.

ஒரு நல்ல அகராதி மூலம் நீங்கள் பின்வருவனவற்றை செய்யலாம்:

- நீங்கள் காணும் அல்லது கேட்கும் அறியாத அல்லது புதிய மொழியின் சொல்லின் பொருளைப் பார்த்து அறியலாம்.
- உங்கள் மொழியில் ஒரு அறியாத மொழியின் சொல்லின் மொழிபெயர்ப்பைக் கண்டறியலாம்.
- ஒரு வார்த்தையின் எழுத்துப்பிழைச் சரிபார்க்கலாம்.
- ஒரு பெயர்ச்சொலின் பன்மை வடிவை அல்லது வேற்றுமை உருபுக்குத் திரிபுற்ற வடிவை மற்றும் வினைச்சொல்லின் முக்கால வடிவங்களை அறிந்துகொள்ளலாம்
- ஒரு சொலைப் பற்றிய பிற இலக்கண தகவல்களைக் கண்டறியலாம்.
- ஒரு சொல்லின் ஒருபொருள்பன்மொழிகளை அல்லது எதிர்ச்சொற்களைக் கண்டறியலாம்.

- ஒரு சொல்லின் சொல்லடிவகைப்பாடுகளைக் கண்டறியலாம்.
- ஒரு சொல்லின் சொல்வகைப்பாட்டைச் சரிபார்க்கலாம்.
- ஒரு சொல்லை எவ்வாறு உச்சரிப்பது என்று கண்டுகொள்ளலாம்.
- ஒரு சொல்லின் நடையியல்சார் வேறுபாட்டை அறிந்துகொள்ளலாம்.
- இயற்கை மொழியில் ஒரு சொல்லைப் பயன்படுத்துவதற்கான எடுத்துக்காட்டுகளைக் கண்டறியலாம்.

ஒரு நல்ல அகராதிப் பயனராக இருக்க, அகராதியை எதற்காகப் பயன்படுத்துவது என்று தெரிந்து கொள்வது போதாது. மேலே பட்டியலிடப்பட்டுள்ள எந்தவொரு நோக்கங்களுக்கும் சிறந்த அகராதி எது என்பதையும் நீங்கள் தீர்மானிக்க வேண்டும். இது போலவே, நீங்கள் தேடுவதை விரைவாக கண்டுபிடிக்க முடியும்; நீங்கள் தேடுவதை நீங்கள் சரியாகக் கண்டுபிடித்தீர்கள் என்பதில் உறுதிபெற வேண்டும்; மற்றும், மிக முக்கியமாக, உங்கள் அகராதியை எப்போது பயன்படுத்த வேண்டும் என்பதை நீங்கள் அறிந்து கொள்ள வேண்டும்.

அகராதி உருவாக்கம்

அகராதி தயாரிக்கும் கலை, அகராதியியல் (லெக்சோகிராபி lexicography), நீண்ட காலமாகக் கவனத்தைப் பெற்றது. அகராதியியல் பொதுவாக ஒரு அகராதியை எழுதுவது அல்லது தொகுப்பது என வரையறுக்கப்படுகிறது. இது பயன்பாட்டு மொழியியலின் ஒரு பகுதியாகும். அகராதியியலின் அடிப்படை அக்கறை 'சொல்' என்பது ஆகும்; இது ஒலியியலியல், உருபனியல், இலக்கணம் மற்றும் நடையியல் என்ற மொழியியலின் வெவ்வேறு கிளைகளில் கையாளப்படுகிறது. 5ஆம் நூற்றாண்டு பி.சி.க்கு முந்தைய காலத்தில் சொற்களஞ்சியத்தில் இந்தியா யஸ்காவின் நிருக்தாவுடன்

(Nirukta of Yaska) ஒரு வளமான பாரம்பரியத்தைக் கொண்டுள்ளது. சொல்சார் தரவு அல்லது அகராதியின் பொருளைக் குறிக்க நிகண்டு என்ற சொல் பயன்படுத்தப்பட்டுள்ளது. அப்போதிருந்து பல அகராதிகள் தயாரிக்கப்பட்டுள்ளன, அவற்றில் பெரும்பாலானவை இன்றும் கிடைக்கின்றன.

இயந்திரத்தால் பயன்படுத்தப்படும் அகராதி வழக்கமான வெளியிடப்பட்ட அகராதிகளிலிருந்து அவை எவ்வாறு ஒழுங்கமைக்கப்படுகின்றன மற்றும் அவை எவ்வாறு தகவல்களைக் குறிக்கின்றன என்பதில் வேறுபடுகின்றன. வழக்கமான அகராதியில் கிடைக்கும் தகவல்கள் வரையறுக்கப்பட்டவை மற்றும் நிலையானவை. வழக்கமான அகராதி அகர வரிசைப்படி சொற்களை ஒழுங்குபடுத்துகிறது. ஒவ்வொரு சொல்லுக்கும் பெயர்ச்சொல், வினைச்சொல், பெயரடைச்சொல், வினையடைச்சொல் போன்ற சொல்வகைப்பாடு தான் அகராதியிலிருந்து ஒரு பயனர் பெறும் முதல் தகவல். ஆக்கப்பட்ட வடிவங்களின் தகவல், உருவாக்கப்பட்ட வடிவங்களின் இலக்கண மற்றும் பொருண்மைசார் தகவல்கள், சொற்றொடர்கள் மற்றும் அவற்றின் நீட்டிப்புகள், எல்லா அளவிலான அர்த்தங்களையும் மேற்கோள் காட்டுதல் போன்ற மொழியின் கிட்டத்தட்ட அனைத்து உள்ளீடுகளையும் பற்றிய முழுமையான தகவல்கள் ஆகஸ்போர்டு ஆங்கில மொழி அகராதி போன்ற விரிவான அகராதியில் மட்டுமே கிடைக்கிறன (தலா 1500 க்கும் மேற்பட்ட பக்கங்களைக் கொண்ட 20 தொகுதிகள்).

முன்னர் கூறியது போல, பல்வேறு நோக்கங்களுக்காகப் பயன்படுத்தப்படும் இயந்திரம் படிக்கக்கூடிய வடிவத்தில் ஒரு அகராதி கொடுக்கப்பட்ட தேவைக்கு தானாகவே தகவல்களைப் பிரித்தெடுக்க உதவுகிறது. தொடர்புடைய தகவலுடன் சொல்சார் தரவு ஒரு கட்டளையிடப்பட்ட கட்டமைப்பில் பயன்படுத்தப்படுகிறது, இது

வெளிப்படையாக நிலையான நடைமுறை மூலம் பிரித்தெடுக்கப்படலாம். தகவல்களைப் பயன்படுத்துவது வழக்கமான அகராதியிலிருந்து வேறுபட்டது. எடுத்துக்காட்டாக, இது அகர வரிசைப்படி இருக்க தேவையில்லை. தேடல் செயல்முறை உருபனியலிருந்து தொடரியல், தொடரியலிருந்து பொருண்மையியல் பாகுபடுத்தலுக்கு நகரும் என்று சொல்சார் தகவல்கள் வரிசைப்படுத்தப்பட்டுள்ளன. ஒரு இயந்திர அகராதி திறந்த-முடிவானது மற்றும் தகவல் சேகரிப்பைக் கட்டுப்படுத்தாது. நவீன தொழில்நுட்பத்தின் நன்மைகளை எடுத்துக் கொண்டால், அகராதியின் அளவை எந்த அளவிலும் அதிகரிக்க முடியும். ஒரு குறிப்பிட்ட சொல்சார் அலகின் தகவலில் அனைத்து ஊடுருவல் மற்றும் வழித்தோன்றல் வடிவங்கள் மற்றும் அவற்றுடன் தொடர்புடைய தகவல்கள், ஒத்த சொற்கள்/ஒருபொருள்பன்மொழிகள் அல்லது எதிர்ச்சொற்கள், தேவையான கலைக்களஞ்சிய அல்லது நடைமுறை அறிவு, வழக்கமான தகவல்களைத் தவிர, இலக்கு மொழிகளுக்கும் பொருந்தும். எந்தவொரு நோக்கத்திற்காகவும் இயந்திரம் அவற்றைப் புரிந்துகொண்டு மீட்டெடுக்கும் வகையில் அகராதியில் உள்ள இந்த தகவல்களை ஒழுங்கமைக்க முடியும்.

கணிப்பொறி வழி அகராதி உருவாக்குவதற்கான தேவை

கணினியைப் பயன்படுத்தி, பலர் உழைப்பால் பல்லாண்டுகாலம் செய்யப்படும் அகராதித் தயாரித்தலை மிகக் குறுகிய காலத்தில் குறைந்த ஆட்களின் சக்தியால் செய்து முடித்துவிடலாம். கணிப்பொறியில் பயன்படுத்தத்தக்க விதத்தில் மின் அகராதிகளையும் தயாரிக்கலாம். இத்தகைய அகராதி கணிப்பொறி வழி மொழி கற்பித்தல், மொழிபெயர்த்தல் போன்ற செயற்பாடுகளுக்குப் பெரிதும் பயன்படும்.

கணினி வழி அகராதி (Computational lexicon)

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

பல துறைகளைப் போல அகராதியியல் துறையிலும் கணினி அகராதியியாரின் உதவியாளரின் உதவிக்கு வந்தது; இதன் காரணமாக நல்ல தரமான மற்றும் நன்கு ஆவணப்படுத்தப்பட்ட அச்சிடப்பட்ட அகராதி உற்பத்தி எளிதாகிவிட்டது. ஆரம்பத்தில் தரவு சேகரிப்பு, டைப் செட்டிங், தரவு மேலாண்மை போன்ற அகராதி முயற்சிகளுக்கு உதவ ஒரு கருவியாக கணினி பயன்படுத்தப்பட்டது. பின்னர், கணினிக்கும் அகராதிக்கும் இடையிலான தொடர்பு பரந்த அளவிலான பகுதிகளில் அவசியமானது. கணினியின் மூலம் மொழிகளைக் கற்பிப்பதற்கும் கற்றுக்கொள்வதற்கும் ஒரு கருவியாக உருவாக்கப்பட்ட எலக்ட்ரானிக்/மின் அகராதி எனப்படும் கணினியில் உள்ள லெக்சிகன், உறவின் அடுத்த உயர் மட்டமாக வந்தது. கணக்கீட்டு அகராதி, இயற்கையான மொழி செயலாக்கம் அல்லது கணக்கீட்டு மொழியியல் துறையில் தவிர்க்க முடியாத பகுதியாகும், அதன் தற்போதைய நிலை.

கணிப்பொறி வழி அகராதி உருவாக்குதல்

மொழி கற்றலில் கணிப்பொறியின் பங்கு மகத்தானது. ஒரு மொழியின் சொல் அதன் இலக்கணக் கூறு, பொருள், பயன்பாடு மற்றும் வாக்கியத்தில் அதன் பயன்பாடு போன்றவைகளை அகராதி விளக்குகிறது. மேலும் ஒரு சொல்லுக்கு நிகரான பிறமொழி சொற்களையும் இருமொழி அகராதிகள் விளங்குகின்றன. அகராதி தயாரிக்கும் பணி எளிதானதல்ல. மிகுந்த நேரமும், அதிக உடல் உழைப்பும் பொருட்செல்வம் ஆகும். ஒரு அகராதியை உருவாக்க நீண்டகாலம் தேவைப்படும். எனவே அகராதி உருவாக்க கணிப்பொறிகள் பயன்படுத்தப்படுகின்றன.

கணிப்பொறி வழி அகராதி: வரலாற்றுச் சுருக்கம்

மாத்தூஸ் அர்னால்ட் (Mathews Arnald) என்ற ஆங்கில கவிஞர் தன் கவிதைகளுக்கான சொல்லடைவை முதன் முதலாக கணிப்பொறி வழி உருவாக்கம் செய்தார். இதற்கு IBM-704 என்ற கணிப்பொறி பயன்படுத்தப்பட்டது. கணிப்பொறி வழி சொல்லடைவினை முதன் முதலாக பேராசிரியர் ஸ்டீபன் (Stephen) உருவாக்கினார். இதனால் தொடர்ந்து Oxford அகராதி உருவாக்க சொல்லடைவு (Concordance) கணிப்பொறி வழி உருவாக்கப்பட்டது. இதைத் தொடர்ந்து ஆங்கிலம் மற்றும் பிற மொழிகளில் அகராதி உருவாக்க கணிப்பொறிகள் ஆங்காங்கே பயன்படுத்தப்பட்டன. தமிழ்ப்பல்கலைக்கழகத்தில் சங்க இலக்கியத்திற்கான அகராதி உருவாக்க கணிப்பொறி பயன்படுத்தப்பட்டது. இதுவே முதல் முயற்சி ஆகும். கிரியா என்னும் தற்கால அகராதி கணிப்பொறி வழி உருவாக்கப்பட்டுள்ளது. இந்த அகராதி அச்ச வடிவிலும், குறுந்தகடுகளிலும் கிடைக்கப் பெறுகின்றன.

அகராதி உருவாக்கும் படிநிலைகள்

பொதுவாக அகராதி ஒன்றை தயாரிப்பதற்கு தேவைப்படும் படிநிலைகள்:

1. செய்திகளை சேகரித்தல், ஆராய்ந்து வகைப்படுத்துதல் அகர வரிசைப்படுத்துதல் என்பது முதற்கட்ட நிலைகளாகும்.
2. சொல்லுக்கான இலக்கணக்கூறு ஆராய்தல், பொருள் எழுதுதல், பிறமொழி சொற்களை எழுதுதல், சொல்லின் உபயோகத்தை விளக்குதல் இவைகளை இரண்டாம் நிலைப் பணியாகும்.

இதில் முதல் கட்ட பணி எளிமையாக கணிப்பொறி படுத்த இயலும். இரண்டாம் கட்டப்பணியை மனித முயற்ச்சியாலேயே செய்ய இயலும்.

கணிப்பொறி வழி அகராதி: முறைமைகள்

கணிப்பொறி வழி அகராதி உருவாக்குவதற்கு கீழ்க்கண்ட முறைமைகள் பின்பற்றப்படுகின்றன.

1. செய்திகளை உள்ளீடு செய்தல்
2. செய்திகளை வார்த்தைகளாகப் பிரித்தல்

இந்நிலையில் ஒரு வாக்கியத்திலிருந்து சொற்கள் தனியாக பிரிக்கப்பட்டு பின்வரும் பட்டியல் உருவாக்கப்படுகிறது.

சொல்	வரி எண்
WORD	LINE NUMBER

1. மேற்கண்ட பட்டியலில் சொற்கள் அவைகள் இடம் பெறும் வரிசையில் இடம் பெறுகின்றன. மேலும் ஒரு சொல் உரையில் பல இடங்களில் பயன்படுத்தப்பட்டிருந்தால் அத்தனை முறை அச்சொல் இப்பட்டியலில் இடம்பெறும்.
2. மேற்கண்ட பட்டியல் அகரவரிசைப் படுத்தப்படுகிறது. இது மேற்கண்ட பட்டியலில் சொற்கள் இடம்பெறும் வரிசை முறையை மாற்றுகிறது. ஒரு சொல் பல முறை இடம் பெற்றிருப்பினும் அது அடுத்தடுத்து வரிசைமுறை பெறுகிறது.
3. இவ்வாறு அகரவரிசைப் படுத்தப் பட்டுள்ள பட்டுள்ள பட்டியலிலிருந்து கீழ்க்கண்ட பட்டியல் உருவாக்கப்படுகிறது.

சொல்	நிகழ்வெண்ணிக்கை	சொல்	இடம்பெறும்
(WORD)	FREQUENCY	இடம்;	
		LIST OF LINE NUMBER	

இப்பட்டியலின் உதவிக் கொண்டு பின்வரும் பட்டியல் உருவாக்கப்படுகின்றது.

சொல் WORD	நிகழ்வெண்ணிக்கை FREQUENCY	சொல் இடம்பெறும் வரிசை எண் LIST OF LINE NUMBER	இடம் பெறும் வரி CITATION
--------------	------------------------------	---	-----------------------------

மேற்கண்ட பட்டியலிலிருந்து ஒரு சொல் உரையில் எத்தனை முறை பயன்படுத்தப்பட்டுள்ளது என்பதை அறியலாம். அச்சொல் இடம்பெறும் வரி தெரிவதால் அதனைக்கொண்டு அச்சொல்லின் இலக்கணக்கூறு பொருள் உபயோகம் போன்றவைகளை ஆராய்ந்து எழுதலாம். இதன்மூலம் விவரங்கள் துல்லியமாகவும் விரைவாகவும் கிடைக்கப்பெறுகின்றன.

அகராதி உருவாக்குவதில் இயற்கை மொழியாய்வு

கணிப்பொறி வழி அகராதி தயாரிக்கப்படும்போது கிடைக்கப்படும் இறுதி படிவம் சொல்லடைவு படிவம் போன்று இருக்கும். இதில் பெயர் சொற்கள் மற்றும் வேற்றுமை உருபுகளுடன் கூடிய பெயர் சொற்கள் மற்றும் வேற்றுமை உருபுகளுடன் கூடிய பெயர் சொற்கள் தனித்தனிச் சொற்களாகவே இடம்பெறும். எடுத்துக்காட்டாக கணிப்பொறி, கணிப்பொறியை, கணிப்பொறியில், கணிப்பொறியால், கணிப்பொறிக்கு போன்ற சொற்கள் தனிச்சொற்களாகவே இடம்பெறும். இவைகளைச் சொல் வடிவங்கள் என மொழியியலாளர்கள் கூறுகின்றனர். ஆனால் அகராதியில் இவைகள் தனிச்சொற்களாக இடம்பெறும் (அல்லது) இடம்பெறாது. கணிப்பொறி என்ற பெயர்ச்சொல் மட்டுமே அகராதியில் இடம்பெறும். இது போன்ற அகராதிச் சொற்களை மொழியியலாளர்கள் சொல்லன்கள்/lexmes என்கின்றனர். இது போன்று

படி என்ற வினைச் சொல்லுக்கு *படித்தான்*, *படித்தார்*, *படித்தது* போன்ற எண்ணற்ற சொல் வடிவங்கள் இருப்பினும் *படி* என்பதே அகராதிச் சொல்லாகும். கணிப்பொறி வழி சொல்லடைவு தயாரிக்கும்போது சொல் வடிவங்களே இடம்பெறும். இதிலிருந்து அகராதிச் சொற்களை மனித முயற்சியினாலேயே மேற்கொள்ள வேண்டும். இப்பணியையும் முழுமையாக கணிப்பொறிப்படுத்த உருபனியல் ஆய்வு மேற்கொள்ள வேண்டும். இவ்வாறாக சொல் வடிவத்திலிருந்து (word forms) உருபனியல் பகுப்பாய்வு மூலம் அகராதி சொற்களை (lexems/dictionary words) உருவாக்கும் முறையை மொழியியலாளர்கள் சொல்லனாக்கம்/lemmatisation என்று அழைக்கின்றனர். இம்முறையை கணிப்பொறிப்படுத்த முயற்சிகள் மேற்கொள்ளப்பட்டு வருகிறது. இதற்கு இயற்கை மொழி ஆய்வு உதவுகிறது.

கணிப்பொறிவழி அகராதி, இயற்கை மொழி ஆய்வுத் (NLP) துறையில் பரவலான பயன்பாடுகளுக்கு ஒரு இயந்திரத்தால் பயன்படுத்தப்படும் ஒரு அகராதி ஆகும். இயந்திரம் படிக்கக்கூடிய அகராதி (Machine Readable Dictionary (MRD/எம்ஆர்டி), மின்னகராதி (எலக்ட்ரானிக் டிசனரி) (Electronic Dictionary (ED/இடி), கணினிசார் அகராதி (கம்ப்யூட்டேஷனல் லெக்சிகன்) (Computational Lexicon (CL/சிஎல்) போன்றவை இதைக் குறிக்க ஒருபொருள்பன்மொழியாகப் பயன்படுத்தப்படும் சொற்கள். எந்தவொரு இயற்கை மொழி ஆய்வு அமைப்பிலும் இது முக்கியமான மற்றும் மிகவும் தேவையான கருவிகளில் ஒன்றாகும். ஒன்று அல்லது வேறு வடிவத்தில், இயந்திர மொழிபெயர்ப்பு (Machine Translation (MT/எம்டி), உரைப் பகுப்பாய்வு மற்றும் உருவாக்கம், மொழி கற்பித்தல் மற்றும் மொழி கற்றல் போன்றவை முதல் அலுவலக தானியக்கம்/ஆட்டோமேஷன் போன்ற கணினி அமைப்புகளின் வளர்ச்சி வரை அனைத்து இயற்கை மொழி அடிப்படையிலான அமைப்புகளுக்கும் இது

ஒரு முன்நிபந்தனை. பேச்சு பகுப்பாய்வு மற்றும் தொகுப்பு, இயற்கையான மொழி இடைமுகங்கள், தகவல் மீட்டெடுப்பு மற்றும் பொதுவாகத் தகவல்தொடர்புகளுக்கு துணைபுரியும் கருவி ஆகும்.

கணினி வாசிக்கக்கூடிய வடிவத்தில் ஒரு அகராதி கிடைப்பது மிகவும் மதிப்புமிக்க வளத்தை வழங்குகிறது; இது உருபனியல் அல்லது தொடரியல் பாகுபடுத்தல், பொருண்மையியல் ஆய்வு, எழுதப்பட்ட மற்றும் பேசும் சொற்களை அறிதல் போன்றவற்றின் மூலம் மேற்கூறிய நோக்கங்களுக்காக நோக்கப்படலாம். இயந்திரம் படிக்கக்கூடிய அகராதி தானாகவே தகவல்களைப் பிரித்தெடுக்க உதவுகிறது. ஒரு அமைப்பு. இங்கே சுதந்திரமாக உள்ளீடுகளின் தொடர்ச்சியான ஒரு கட்டமைப்பில் சொல்சார்/லெக்சிக்கல் சம்பந்தப்பட்ட தகவல்கள் பயன்படுத்தப்படுகின்றன; எனவே உத்தரவிடப்பட்டு, எந்தவொரு உள்ளீட்டையும் வெளிப்படையாக நிலையான நடைமுறை மூலம் காணலாம். ஒரு இயந்திர அகராதி தானியங்கி மொழிபெயர்ப்பிற்கான ஒரு கருவியாக செயல்படுகிறது மற்றும் இலக்கு மொழியைக் கற்க உதவுகிறது. இந்த அனைத்து நோக்கங்களுக்காகவும் அகராதி பல்வேறு நிலைகளில் தகவல்களைக் கொண்டிருக்க வேண்டும் - ஒலியனியல், உருபனியல், தொடரியல், பொருண்மையியல் மற்றும் பயன்வழியியல். இந்த அனைத்து தேவைகளையும் பூர்த்தி செய்ய ஒரு அகராதியை உருவாக்குவது எந்த இயற்கைமொழி ஆய்வு (என்.எல்.பி) அமைப்பிலும் முதல் மற்றும் முக்கியமான படியாகும். நவீன இலக்கணக் கோட்பாடுகளின்படி, கருத்தாடல் பகுப்பாய்விற்கான தகவல்களின் முக்கிய ஆதாரமாக அகராதி உள்ளது. முழுமையாக பிரதிநிதித்துவப்படுத்தப்பட்ட அகராதி தயாராக இருந்தால், இயற்கையான மொழி அடிப்படையிலான அமைப்பின் வளர்ச்சி பாதி வழிக்கு மேல் எனக் கருதப்படுகிறது. மேற்கூறிய உண்மைகளின் வெளிச்சத்தில், ஒரு அகராதியை உருவாக்குவதற்கான

முன்மொழியப்பட்ட பணியில், தரவுகளுக்கான ஒரு கட்டமைப்பு மட்டுமே முயற்சிக்கப்படுகிறது, ஏனெனில் ஒரு முழு அகராதி அமைப்பது வெளிப்படையாக ஒரு குழு மற்றும் கணக்கீட்டு குழுவின் பணியாகும் மொழியியலாளர்கள். கணினி விஞ்ஞானிகளின் உதவியுடன் கணக்கீட்டு செயல்படுத்தல் மற்றும் கட்டடக்கலை ஆதரவுகள் செய்யப்படுகின்றன.

அகராதி தயாரிப்பதற்கான சில மென்பொருள்கள்

அகராதி தயாரிப்பதற்கான சில மென்பொருள்கள் வெளிவந்துள்ளன. (எடுத்துக்காட்டாக Tool Box என்ற மென்பொருள் அகராதி தயாரிக்கப் பெரிதும் உதவும்.) அவற்றைப் பயன்படுத்தி அகராதிகள் தயாரிக்கலாம். தொடக்க காலத்தில் DATA BASE, DATA BASE MANAGEMENT SYSTEMSவ் ஃபாக்ஸ்புரோ, விஷுவல் ஃபாக்ஸ்புரோ போன்றவை அகராதி தயாரித்தலுக்குப் பெரிதும் பயன்பட்டன. MySQL, MS Excel Worksheet என்பன அகராதி ஆக்கத்தை எளிதாக்கும்.

அகராதியின் ஒவ்வொரு சொற்பதிவிலும் தலைப்புச் சொல், எழுத்துப்பெயர்ப்பு, சொல் வகைப்பாடு, பல்பொருண்மைகள் போன்ற தகவல்கள் அடங்க வேண்டும் என்பதன் அடிப்படையில் தரவு மையத்தை உருவாக்க வேண்டும். கீழ்வரும் எடுத்துக்காட்டில் முக்காற்புள்ளிக்கு இடப் பக்கம் வருவனவற்றைப் பீல்டுகளின் பெயர்களாகவும் முக்காற்புள்ளிக்கு வலப்பக்கம் வருவனவற்றை அவற்றில் நிரப்பப்பட வேண்டிய செய்திகளாகவும் எடுத்துக்கொள்ளலாம். ஒவ்வொரு சொற்பதிவும் ஒரு பதிவாக அமையும்.

தலைப்புச் சொல்: எவு

எழுத்துப்பெயர்ப்பு: eevu

சொல் வகைப்பாடு: செ.பொ.குன்றா.வி. (செயப்படுபொருள் குன்றா வினை)

பொருள் 1: தூண்டுதல்

எடுத்துக்காட்டு 1: குண்டர்களை ஏவி எங்களைப் பயமுறுத்தினான்.

பொருள் 2 : 2. கட்டவிழ்த்துவிடல்

எடுத்துக்காட்டு 2 : அடக்கு முறையை ஏவிவிட்டு மக்களின் போராட்டத்தை அடக்கி விட இயலாது.

பொருள் 3 : அனுப்புதல், செலுத்துதல்

எடுத்துக்காட்டு 3 : ஏவுகணைகளை ஏவிப் பலத்த சேதம் விளைவித்தனர்.

வழியமைப்புகளைப் பயன்படுத்திச் சொற்பதிவைப் பின்வருமாறு அகராதி வடிவில் தரலாம்.

ஏவு. eevu செ. பொ. குன்ற.வி. 1. தூண்டுதல்; குண்டர்களை ஏவி எங்களைப் பயமுறுத்தினான். 2. கட்டவிழ்த்துவிடல்; அடக்கு முறையை ஏவிவிட்டு மக்களின் போராட்டத்தை அடக்கிவிட இயலாது. 3. அனுப்புதல், செலுத்துதல்; ஏவுகணைகளை ஏவிப்பலத்த சேதம் விளைவித்தனர்.

அகராதி தொகுப்பின் சில அடிப்படை சிக்கல்கள்.

அகராதித் தொகுத்தல் அறிவியலான அகராதியியல் சொல்லியியலுடன் நெருக்கமாக இணைக்கப்பட்டுள்ளது; இவை இரண்டும் ஒரே சிக்கல்களைக் கையாளுகின்றன: சொற்களின் வடிவம், சொற்களின் பொருள், சொற்களின் பயன்பாடு மற்றும் சொற்றொகை அலகுகளின் தோற்றம். மேலும் அவை ஒன்றுக்கொன்றான சாதனைகளைப் பயன்படுத்துகின்றன

அகராதி-தொகுப்பின் சில அடிப்படை சிக்கல்கள்:

1) சேர்ப்பதற்கான சொல்சார் அலகுகளின் தேர்வு,

2) அவற்றின் வரிசை அமைப்பு,

- 3) பதிவுகளின் வரிசை அமைப்பு,
- 4) சொல்-அர்த்தங்களின் தேர்வு மற்றும் வரிசை அமைப்பு (தொகுத்தல்),
- 5) அர்த்தங்களின் வரையறை (the definition of meanings),
- 6) விளக்கப் பொருள் (எடுத்துக்காட்டுகள்) (illustrative material),
- 7) துணைப் பொருள் (supplementary material).

1) சேர்ப்பதற்கான சொல்சார் அலகுகளின் தேர்வு

பின்வருவன தீர்மானிக்க வேண்டியது அவசியம்: அ) சேர்ப்பதற்கு எந்த வகையான சொல்சார் அலகுகள் தேர்வு செய்யப்பவேண்டும்; ஆ) சொற்களின் எண்ணிக்கை; இ) அகராதியில் எதைத் தேர்ந்தெடுக்க வேண்டும், எதை விட்டுவிட வேண்டும்; ஈ) எந்த மொழியின் பேச்சு அல்லது எழுத்து வடிவமா அல்லது இரண்டுமே அகராதி பிரதிபலிக்க வேண்டுமா? உ) அகராதியில் வழக்கற்றுப்போன அலகுகள், தொழில்நுட்பச் சொற்கள், கிளைமொழிச் சொற்கள், பேச்சுவழக்குச் சொற்கள் மற்றும் பலவற்றைக் கொண்டிருக்க வேண்டுமா?

அகராதி எந்த வகையைச் சேர்ந்தது, தொகுப்பாளர்கள் தொடரும் நோக்கம், அகராதியின் வருங்கால பயனர், அதன் அளவு, அகராதி தயாரிப்பாளர்களின் மொழியியல் கருத்துக்கள் மற்றும் வேறு சில கருத்தாய்வுகளைப் பொறுத்து தேர்வு செய்யப்படுகிறது.

2) பதிவுகளின் வரிசை அமைப்பு

பதிவுகளை வழங்குவதில் இரண்டு முறைகள் உள்ளன: அகரவரிசை மற்றும் கிளஸ்டர் வகை (கூடுகளில் வரிசை செய்யப்பட்டுள்ளன, சில கொள்கையின் அடிப்படையில் - ஒரே மூலத்தின் சொற்கள்).

3) பதிவுகளின் கட்டமைப்பு

பல்வேறு வகையான அகராதிகள் அவற்றின் நோக்கம், அவை வழங்கும் தகவல்கள், அவற்றின் அளவு போன்றவற்றில் வேறுபடுவதால், அவை பதிவின் கட்டமைப்பு மற்றும் உள்ளடக்கத்தில் அவசியமான நிலையில் வேறுபடுகின்றன.

மிகவும் சிக்கலான வகை பதிவு என்பது ஒருகால வகையின் பொதுவான விளக்க அகராதிகளில் காணப்படுகிறது (பதிவு பொதுவாகப் பின்வரும் தரவை அளிக்கிறது: ஏற்றுக்கொள்ளப்பட்ட எழுத்துக்கூட்டல் மற்றும் உச்சரிப்பு; ஒவ்வொரு பதிவுச் சொல்லின் சொல்வகைப்பாட்டை குறிப்பது உள்ளிட்ட இலக்கணப் பண்புகள், பெயர்ச்சொற்கள் எண்ணக்கூடியவை எண்ண இயலாதது என்பது, வினைச்சொற்களின் செயப்படுபொருள் குன்றிய அல்லது குன்றாப் பண்பு மற்றும் அவற்றின் ஒழுங்கற்ற இலக்கண வடிவங்கள்; அர்த்தங்களின் வரையறைகள்; நவீனப் பயன்பாடு; விளக்க எடுத்துக்காட்டுகள்; ஆக்கங்கள்; சொற்றொடர்; சொற்பிறப்பியல்; சில சமயங்களில் ஒருபொருள்பன்மொழிகள் மற்றும் எதிர் மொழிகள்.

4) சொல்-அர்த்தங்களின் தேர்வு மற்றும் வரிசை அமைப்பு (தொகுத்தல்).

ஒரு சொல் கொடுக்கப்பட்ட அர்த்தங்களின் எண்ணிக்கை மற்றும் ஒரு அகராதியில் அவற்றின் தேர்வு முக்கியமாக இரண்டு காரணிகளைச் சார்ந்துள்ளது: 1) தொகுப்பாளர்கள் தங்களை எந்த நோக்கத்தில் அமைத்துக்கொள்கிறார்கள் மற்றும் 2) எந்த அளவிற்கு வழக்கற்று, பழமையான, கிளைமொழி அல்லது மிகவும் சிறப்பு வாய்ந்த அர்த்தங்கள் பதிவு செய்யப்பட வேண்டும், பல்பொருள் ஒருமொழியம் மற்றும் ஒப்புருமொழியம் இவற்றில் சிக்கல் எவ்வாறு தீர்க்கப்படுகிறது, சொல்வகைப்பாட்டு மாற்றம் எவ்வாறு நடத்தப்படுகின்றன, ஒரு பல்பொருள் ஒருமொழியின் வெவ்வேறு அர்த்தங்களின் பிரிவு எவ்வாறு செய்யப்படுகிறது போன்றவை. சொல் அர்த்தங்கள் ஒழுங்கமைக்கப்படுவதற்கு குறைந்தது மூன்று வெவ்வேறு வழிகள் உள்ளன: அ) அவற்றின் வரலாற்று வளர்ச்சியின் வரிசையில் (வரலாற்று ஒழுங்கு என

அழைக்கப்படுகிறது), ஆ) பயன்பாட்டின் அதிர்வெண்ணுடன் இணக்கமாக, இது மிகவும் பொதுவான அர்த்தத்துடன் முதலில் (அனுபவ அல்லது உண்மையான) ஒழுங்கு), இ) அவற்றின் தருக்கத் தொடர்பில் (தருக்க வரிசை).

5) அர்த்தங்களின் வரையறை

சொற்களின் அர்த்தங்கள் வெவ்வேறு வழிகளில் வரையறுக்கப்படலாம்: 1) சொற்களைப் பேச்சுப் பொருளாக மட்டுமே கருதும் மொழியியல் வரையறைகள் மூலம், 2) கலைக்களஞ்சிய/என்சைக்ளோபீடிக் வரையறைகள் மூலம், 3) ஒத்த சொற்கள் மற்றும் வெளிப்பாடுகளின் வழிமுறையாக, 4) குறுக்கு குறிப்புகள் மூலம். தேர்வு என்பது சொல்லின் தன்மையைப் பொறுத்தது (பேச்சின் பகுதி, அகராதியின் நோக்கம் மற்றும் அளவு).

6) விளக்கப் பொருள்

இது அகராதியின் வகையைப் பொறுத்தது மற்றும் தொகுப்பாளர்கள் தங்களை அமைத்துக் கொள்ளும் நோக்கத்தைப் பொறுத்தது.

கணினி அகராதி மின்னணு (மின்னணு முறையில் படிக்கக்கூடிய/இயந்திரம் படிக்கக்கூடிய) அகராதிகளின் வடிவமைப்பு, தொகுப்பு, பயன்பாடு மற்றும் மதிப்பீடு ஆகியவற்றைக் கையாள்கிறது. மின்னணு அகராதிகள் வழக்கமான காகித நூல் அமைப்புகளிலிருந்து வடிவம், உள்ளடக்கம் மற்றும் செயல்பாட்டின் அடிப்படையில் வேறுபடுகின்றன. மிக முக்கியமான வேறுபாடுகளில் சில:

- 1) பல்லுடக (மல்டிமீடியா/multimedia) வழிமுறைகளின் பயன்பாடு;
- 2) சாளரங்கள் சார்ந்த மென்பொருளில் செல்லக்கூடிய உதவி குறியீடுகள் (indices);
- 3) ஒலி, அனிமேஷன், ஆடியோ மற்றும் காட்சி (படங்கள், வீடியோக்கள்) கூறுகள் மற்றும் ஊடாடும் பயிற்சிகள் மற்றும் விளையாட்டுகளின் பயன்பாடு;

- 4) பல வழிகளில் வெளியீட்டைக் குறிப்பிட பயனரை அனுமதிக்கும் தேடல் மற்றும் அணுகல் முறைகளின் மாறுபட்ட சாத்தியங்கள்;
- 5) தகவல்களை அணுகுவதும் மீட்டெடுப்பதும் இனி உள், பாரம்பரியமாக அகரவரிசை, அகராதியின் அமைப்பால் தீர்மானிக்கப்படுவதில்லை, ஆனால் உரையின் நேரியல் அல்லாத அமைப்பு;
- 6) ஹைப்பர்லிங்க்களின் பயன்பாடு எளிதாகவும் விரைவாகவும் ஒரு பதிவில் உள்ள சொற்களைக் குறிக்க அல்லது இந்த பதிவுடன் இணைக்கப்பட்ட பிற சொற்களைக் குறிக்க அனுமதிக்கிறது.

மின் அகராதிகளைப் பொறுத்தவரையில், அகராதி கட்டமைப்பின் மூலம் முன்னரே தீர்மானிக்கப்பட்ட பாதையைப் பின்பற்றுவதற்கும், இணைப்புகளின் தேர்வு வழியாக உள்ளீடுகளுக்குள்ளும் உள்ளேயும் உள்ள உறவுகளை வழிநடத்துவதற்கும் முக்கியத்துவம் குறைவாக இருப்பதால் பயனரின் கோரிக்கைகள் அதிகமாகின்றன. எனவே ஒரு மின்னணு அகராதியைப் பயன்படுத்துவதற்கு முன்பு, 'வழக்கமான அகராதித் திறன்களை' தவிர சில ஊடுருவல் மற்றும் தேடல் திறன்களைப் பெறுவது அவசியம்.

நவீன அகராதியியல்

நவீன அகராதியியல் துறையானது அனைத்து வகையான அகராதிகளையும் ஒரு பெரிய எண்ணிக்கையையும் பல்வேறு வகையான அகராதிகளையும் வழங்குகிறது. தமிழ் அகராதிகள் ஒருமொழி அகராதிகள், இருமொழி அகராதிகள், பொது அகராதிகள், சிறப்பு அகராதிகள், சொற்பிறப்பியல் அகராதிகள் மற்றும் தற்காலத் தமிழ் அகராதிகள் எனப் பலவகைப்படும். மேலும் நடைவேறுபாடுகள், கிளைமொழிகள் மற்றும் கொச்சைச்சொற்கள் ஆகியற்றைக் கையாளும் அகராதிகளும் உள்ளன. நவீன அகராதியியல் வரலாற்று மற்றும் பயன்வழியியல் (நடைமுறை) சார்ந்த அல்லது கற்பவரின் அகராதிகளுக்கு இடையில் வேறுபாடு பாராட்டுகிறது.

நடைமுறையில்/பயன்வழியியல் (pragmatics) சார்ந்த அகராதிகள் இலக்கியப் படைப்புகளில் பதிவுசெய்யப்பட்ட சொற்களின் அர்த்தங்களுடன் பக்கவாட்டாக செயல்படும் முக்கிய அர்த்தங்களைப் பதிவுசெய்கின்றன, இதனால் இந்த சொல் உண்மையில் பேச்சில் எவ்வாறு பயன்படுத்தப்படுகிறது என்பது பற்றிய தெளிவான கருத்தை வாசகர்களுக்கு அளிக்கிறது.

அகராதியியலின் நவீன போக்குகள் தரவுத்தொகுதி மொழியியல் (Corpus Linguistics) மற்றும் கணினி மொழியியல் (Computational Linguistics) போன்ற மொழியியலின் கிளைகளின் தோற்றம் மற்றும் விரைவான வளர்ச்சியுடன் இணைக்கப்பட்டுள்ளன.

தரவுத்தொகுதி அடிப்படையிலான அகராதியியல்

தரவுத்தொகுதி அடிப்படையிலான மொழியியல் முக்கியமாக ஒலியியல், ஒலியியலியல், இலக்கணம், நடையியல், கருத்தாடல், அகராதி பல போன்ற மொழியியல் துறைகளில் ஆய்வுகளை நடத்துவதற்காக பல்வேறு மின்னணு தரவுத்தொகுதிகளைத் தொகுப்பதில் ஈடுபட்டுள்ளது. தரவுத்தொகுதி பெரிய மற்றும் முறையான முனைப்புகள் அகும்: அவற்றில் உரையாடல்கள், பத்திரிகை கட்டுரைகள், செய்தித்தாள்கள், விரிவுரைகள், நாவல்களின் அத்தியாயங்கள், பிரசுரங்கள் போன்றவை அடங்கியுள்ளன. தரவுத்தொகுதிகள் மொழி செயல்படும் முறை பற்றிய வளங்களையும் சொல்வளங்களையும் ஆய்வாளர்களுக்கு வழங்குகின்றன.

சொற்கள், தொடரடைவுகள், எழுத்துப்பிழைகள், உச்சரிப்புகள் மற்றும் இலக்கண நிர்மாணங்கள் போன்ற மொழியியல் பண்புக்கூறுகளின் அதிர்வெண், விநியோகம் மற்றும் தனித்துவம் பற்றிய சிறந்த தகவல்களை ஒரு பெரிய மற்றும் நன்கு கட்டமைக்கப்பட்ட தரவுத்தொகுதி வழங்குகிறது. தரவுத்தொகுதி மொழியியலின் வளர்ச்சி தரவுத்தொகுதி அடிப்படையாகக் கொண்ட அகராதியியல் வளர்ச்சிக்கு

உதவுவதோடு புதிய தரவுத்தொகுதி அடிப்படையிலான அகராதிகளை உருவாக்கப் பெரிதும் உதவுகின்றது.

அகராதி தயாரிக்கும் நடைமுறைகளில் தரவுத்தொகுதியின் பயன்பாடு ஒரு அகராதியியலாருக்கு நிறைய வாய்ப்புகளைத் தருகிறது; மிக முக்கியமான வாய்ப்புகளாகப் பின்வருவனவற்றைப் பட்டியலிடலாம்:

- 1) அகராதிகளை மிக விரைவாக தயாரித்து திருத்துதல்; இதனால் மொழி குறித்த புதுப்பித்த தகவல்களை வழங்குதல்;
- 2) அதிக எண்ணிக்கையிலான இயற்கையான எடுத்துக்காட்டுகள் ஆராயப்படுவதால் இன்னும் முழுமையான மற்றும் துல்லியமான வரையறைகளை வழங்குவது;
- 3) மொழியில் நுழையும் புதிய சொற்களின் மேல் அல்லது தற்போதைய சொற்களை அவற்றின் அர்த்தங்களை மாற்றுவது;
- 4) குறிப்பிட்ட வகைகள் மற்றும் வகைகளின் பொதுவான சொற்கள் அல்லது சொற்றொடர்களின் பயன்பாடுகளை விவரிப்பது;
- 5) தரவுத்தொகுதியிலிருந்து பிரித்தெடுக்கப்பட்ட எடுத்துக்காட்டுகளைப் பகுப்பாய்விற்காக மிகவும் அர்த்தமுள்ள குழுக்களாக ஒழுங்கமைத்து, அவற்றின் சொல்லடிவகைப்பாட்டில் சிறப்பு அழுத்தத்தை அளிப்பதை விவரிப்பது / வழங்குவது.
- 6) பரஸ்பர தகவல் கருவிகளின் இருப்பு காரணமாக இணைந்த சொற்களுக்கு இடையில் உறவை நிலைநிறுத்துவதன் காரணமாக சொற்களைக் காட்டிலும் சொல்-சேர்க்கைகளை அழைக்கும் திறன் காரணமாக முன்னர் இருந்ததை விட சொற்றொடர்களையும் சொல்லடி வகைப்பாடுகளையும் முறையாக நடத்துவது;
- 7) ஒரு மொழி கொண்டிருக்கும் கலாச்சார அர்த்தங்களையும் அடிப்படை சித்தாந்தங்களையும் பதிவு செய்தல்.

சில அகராதியியல் முயற்சிகள் சொந்த மின்னணு உரைக் காப்பகங்களைக் கொண்டுள்ளன; அவை தொகுக்கப்பட்ட அகராதியின் வகையைப் பொறுத்துப் பயன்படுத்தப்படுகின்றன.

மின்னணு அகராதி

மின்னணு அகராதி என்பது ஒரு அகராதியாகும், அதன் தரவு டிஜிட்டல் வடிவத்தில் உள்ளது மற்றும் பல்வேறு ஊடகங்கள் மூலம் அணுக முடியும். டேப்லெட் அல்லது டெஸ்க்டாப் கணினிகளில் நிறுவப்பட்ட மென்பொருள், மொபைல் பயன்பாடுகள், இணையப் பயன்பாடுகள் மற்றும் மின்-வாசகர்களின் உள்ளமைக்கப்பட்ட செயல்பாடு உட்பட பல வடிவங்களில் மின்னணு அகராதிகளைக் காணலாம். அவை இலவசமாக இருக்கலாம் அல்லது கட்டணம் செலுத்த வேண்டியிருக்கலாம்.

ஆரம்பகால மின்னணு அகராதிகளில் பெரும்பாலானவை டிஜிட்டல் வடிவில் கிடைக்கப்பெற்ற அச்சு அகராதிகள் ஆகும்: உள்ளடக்கம் ஒரே மாதிரியாக இருந்தது, ஆனால் மின்னணு பதிப்புகள் பயனர்களுக்கு அதிக சக்திவாய்ந்த தேடல் செயல்பாடுகளை வழங்கின. ஆனால் விரைவில் டிஜிட்டல் மீடியாவின் வாய்ப்புகள் பயன்படுத்தப்பட ஆரம்பித்தன. இரண்டு வெளிப்படையான நன்மைகள் என்னவென்றால், இடத்தின் வரம்புகள் (மற்றும் அதன் பயன்பாட்டை மேம்படுத்த வேண்டிய அவசியம்) குறைவான அழுத்தம், எனவே கூடுதல் உள்ளடக்கத்தை வழங்க முடியும்; மற்றும் ஆடியோ உச்சரிப்புகள் மற்றும் வீடியோ கிளிப்புகள் போன்ற பல்லாட்க உள்ளடக்கத்தை உள்ளடக்கும் சாத்தியம் எழுகிறது.

மின்னணு அகராதித் தரவுத்தளங்கள் (Electronic dictionary databases), குறிப்பாக மென்பொருள் அகராதிகளுடன் சேர்க்கப்பட்டுள்ளன: பெரும்பாலும் விரிவானவை மற்றும் 500,000 தலைப்புச் சொற்கள் மற்றும் வரையறைவிளக்கங்கள், வினைச்சொல் திரிபு அட்டவணைகள் மற்றும் இலக்கணக் குறிப்புப் பிரிவு

ஆகியவற்றைக் கொண்டிருக்கலாம். இருமொழி மின் அகராதிகள் மற்றும் திரிபு மொழிகளின் ஒருமொழி அகராதிகள் பெரும்பாலும் ஒரு ஊடாடும் வினைத்திரிபாக்கியை (interactive verb conjugator) இணைக்கும், மேலும் அவை பகுதியாக்கம் (word stemming) மற்றும் சொல்லனாக்கம் (lemmatization) செய்யும் திறன் கொண்டவை.

மின்னணு அகராதிகளின் வெளியீட்டாளர்கள் மற்றும் உருவாக்குபவர்கள் தங்கள் சொந்த அகராதியிடமிருந்து சொந்த உள்ளடக்கத்தை வழங்கலாம், அச்ச வெளியீடுகளிலிருந்து உரிமம் பெற்ற தரவு அல்லது இரண்டையும் வழங்கலாம்: பாபிலோன் (Babylon) மெரியம் வெப்ஸ்டரிடமிருந்து பிரிமியம் உள்ளடக்கத்தை வழங்குகிறது, மற்றும் அல்ட்ராலிங்குவா (Ultralingua) காலின்ஸ் (Collins), மாசன் (Masson) மற்றும் சைமன் & ஸ்கஸ்டர் (Simon & Schuster) ஆகிவற்றிலிருந்து கூடுதல் பிரிமியம் உள்ளடக்கத்தை வழங்குகிறது, மற்றும் பேரகான் மென்பொருள் (Paragon Software) டூடென் (Duden), பிரிட்டானிக்கா (Britannica), ஹார்ப் (Harrap), மரியம் வெப்ஸ்டர் (Merriam-Webster) மற்றும் ஆக்ஸ்போடிலிருந்து (Oxford) அசல் உள்ளடக்கத்தை வழங்குகிறது.

மின்னணு அகராதிகள் லத்தீன் எழுத்துக்களுடன், அரபு, பாரசீகம், சீனம், தேவநாகரி, கிரேக்கம், ஹீப்ரு, ஜப்பானிய, கொரியன், சிரிலிக் மற்றும் தாய் உள்ளிட்ட லோகோகிராஃபிக் மற்றும் வலமிருந்து இட எழுத்துகளிலும் கிடைக்கின்றன.

அகராதி மென்பொருள் (Dictionary software)

அகராதி மென்பொருள் பொதுவாக கையில் வைத்திருக்கும் அகராதிகளின் நோக்கத்தை விட அதிகமாக உள்ளது. Langenscheidt, Collins-Reverso, Oxford University Press, Duden, American Heritage மற்றும் Hachette போன்ற பாரம்பரிய அச்சிடப்பட்ட அகராதிகளின் பல வெளியீட்டாளர்கள் டெஸ்க்டாப் மற்றும் லேப்டாப்

கணினிகளில் பயன்படுத்த தங்கள் ஆதாரங்களை வழங்குகிறார்கள். இந்த நிரல்களைப் (programs) பதிவிறக்கம் செய்யலாம் அல்லது CD-ROM-இல் வாங்கி நிறுவலாம். iFinger, ABBYY Lingvo, Collins-Ultralingua, மொபல் ஒழுங்குமுறைகள் (Mobile Systems) மற்றும் பேரகான் மென்பொருள் (Paragon Software) போன்ற சிறப்பு மின்னணு அகராதி வெளியீட்டாளர்களிடமிருந்து பிற அகராதி மென்பொருள்கள் கிடைக்கின்றன. சில மின்னணு அகராதிகள் மென்பொருள் உருவாக்குநர்கள் மற்றும் அகராதியாளர்களால் நிர்வகிக்கப்படும் ஆன்லைன் விவாத மன்றத்தை வழங்குகின்றன.

கையடக்க அகராதிகள் அல்லது PEDகள்

கையடக்க எலக்ட்ரானிக் அகராதிகள் (Handheld electronic dictionaries), "பாக்கெட் எலக்ட்ரானிக் அகராதிகள்" ("pocket electronic dictionaries அல்லது PED) என்றும் அழைக்கப்படுகின்றன, முழு விசைப்பலகைகள் மற்றும் LCD திரைகளுடன் முழுமையான சிறிய கிளாம்புஷல் லேப்டாப் கணினிகளை (clamshell laptop computers) ஒத்திருக்கும். அவை முழுவதுமாக எடுத்துச் செல்லக்கூடியவையாக இருப்பதால், அகராதிகள் பேட்டரி மூலம் இயங்கும் மற்றும் நீடித்த உறைப் பொருட்களால் செய்யப்பட்டவை. உலகம் முழுவதும் தயாரிக்கப்பட்டாலும், கையடக்க அகராதிகள் குறிப்பாக ஜப்பான், கொரியா, தைவான், சீனா மற்றும் அண்டை நாடுகளில் பிரபலமாக உள்ளன, அங்கு ஆங்கிலத்தை இரண்டாம் மொழியாகக் கற்கும் பல பயனர்களுக்கு அவை தேர்வு அகராதியாகும். கையடக்க அகராதிகளின் சில அம்சங்களில் ஸ்ட்ரோக் ஆர்டர் அனிமேஷன் (stroke order animations), குரல் வெளியீடு (voice output), கையெழுத்து அறிதல் (handwriting recognition), மொழி-கற்றல் திட்டங்கள், கணிப்பான் (calculator), தனிப்பட்ட டிஜிட்டல் உதவியாளர் (personal digital assistant/PDA) போன்ற அமைப்பாளர் செயல்பாடுகள், நேர

மண்டலம் (time zone) மற்றும் நாணய மாற்றிகள் (currency converters) மற்றும் குறுக்கெழுத்து புதிர்கள் தீர்வி(crossword puzzle solvers) ஆகியவை அடங்கும். பல மொழிகளுக்கான தரவைக் கொண்ட அகராதிகளில் "ஜம்ப்" அல்லது "தவிர்-தேடல்" அம்சம் இருக்கலாம், இது பயனர்கள் சொற்களைத் தேடும் போது அகராதிகளுக்கு இடையில் செல்ல அனுமதிக்கிறது, மேலும் முடிவுகளில் காட்டப்படும் சொற்களை மேலும் பார்க்க அனுமதிக்கும் தலைகீழ் மொழிபெயர்ப்பு நடவடிக்கை. பல உற்பத்தியாளர்கள் உரிமம் பெற்ற அகராதி உள்ளடக்கத்தைப் பயன்படுத்தும் கையடக்க அகராதிகளை உருவாக்குகிறார்கள் அவை மெரியம் வெப்ஸ்டர் அகராதி மற்றும் தெசரஸ்/சொற்களஞ்சியம் போன்ற தரவுத்தளத்தைப் பயன்படுத்துகின்றன, மற்றவர்கள் தங்கள் சொந்த அகராதியாளர்களிடமிருந்து தனியுரிம தரவுத்தளத்தைப் பயன்படுத்தலாம். பயனர்கள் தங்கள் கையடக்க அகராதிகளில் மெமரி கார்டுகள் (விரிவாக்கக்கூடிய மற்றும் அர்ப்பணிப்பு இரண்டும்), CD-ROM தரவு மற்றும் இணையப் பதிவிறக்கங்கள் மூலம் உள்ளடக்கத்தைச் சேர்க்கலாம். உற்பத்தியாளர்களில் AlfaLink, Atree, Besta, Casio, Canon, Instant Dict, Ectaco, Franklin, Iriver, Lingo, Maliang Cyber Technology, Compagnia Lingua Ltd., Nurian, Seiko மற்றும் Sharp ஆகியவை அடங்கும்.

மொபைல் சாதனங்களில் அகராதிகள்

அனைத்து வகையான அகராதிகளும் ஸ்மார்ட்போன்களுக்கான பயன்பாடுகளாகவும், ஆப்பிள் ஐபாட், பிளாக்பெர்ரி பிளேபுக் (BlackBerry PlayBook) மற்றும் மோட்டோரோலா ஜூம் (Motorola Xoom) போன்ற டேப்லெட் கணினிகளுக்காகவும் கிடைக்கின்றன. மொழிபெயர்ப்பாளர்கள் மற்றும் மொழி கற்பவர்களின் தேவைகள், பல மொழி இணைகளுக்கான இருமொழி அகராதிகளுக்கான மற்றும் லாங்மேன் டிசுஷனரி ஆஃப் கன்டெம்பரரி இங்கிலீஷ் மற்றும் மேக்மில்லன் ஆங்கில அகராதி போன்ற பெரும்பாலான நன்கு அறியப்பட்ட

ஒருமொழி கற்பவர்களின் அகராதிகளுக்குகான செயலிகளுடன் (apps) சிறப்பாகப் பூர்த்தி செய்யப்படுகின்றன.

ஆன்லைன் அகராதிகள்

பல வகையான ஆன்லைன் அகராதிகள் உள்ளன, பின்வருவன இதில் அடங்கும்: பல்வேறு நோக்கீட்டு வெளியீட்டாளர்களிடமிருந்து உரிமம் பெற்ற தரவுகளுக்கான அணுகலை வழங்கும் திரட்டி தளங்கள் (Aggregator sites). அவை பொதுவாக ஒருமொழி மற்றும் இருமொழி அகராதிகள், ஒன்று அல்லது அதற்கு மேற்பட்ட சொற்களஞ்சியம் மற்றும் தொழில்நுட்ப அல்லது சிறப்பு அகராதிகளை வழங்குகின்றன. (எ.கா. TheFreeDictionary.com, Dictionary.com, Kotobank மற்றும் Goo அகராதி)

'பிரீமியம்' அகராதிகள் சந்தாவில் கிடைக்கும் (எ.கா. ஆக்ஸ்போர்டு ஆங்கில அகராதி, வான் டேல் மற்றும் கென்கியூஷா ஆன்லைன் அகராதி)

ஒற்றை வெளியீட்டாளரின் அகராதிகள், பயனருக்கு இலவசம் மற்றும் விளம்பரத்தால் ஆதரிக்கப்படும் (எ.கா. காலின்ஸ் ஆன்லைன் அகராதி, டூன் ஆன்லைன், லாரூஸ் இருமொழி அகராதிகள், மேக்மில்லன் ஆங்கில அகராதி மற்றும் மெரியம்-வெப்ஸ்டர் கற்றோர் அகராதி)

அல்ஜிமீன் நெடர்லாண்ட்ஸ் வூர்டன்போக் [Algemeen Nederlands Woordenboek [nl] (என்எல்) (ANW), (ஏஎன்டபிள்யூ), மற்றும் டென் டான்ஸ்கே ஆர்ட்போக் (Den Danske Ordbog) [da/டா] போன்ற வணிக ரீதியற்ற வெளியீட்டாளர்களிடமிருந்து (பெரும்பாலும் அரசாங்க நிதியுதவி பெறும் நிறுவனங்கள்) அகராதிகள் இலவசமாகக் கிடைக்கின்றன.

ஆன்லைன் அகராதிகள் தொடர்ந்து புதுப்பிக்கப்பட்டு, மொழி மாற்றத்தைத் தெரிந்துகொள்ளும். பல புதிய சொற்களில் வலைப்பதிவுகள் மற்றும் அம்சங்கள் போன்ற கூடுதல் உள்ளடக்கம் உள்ளது. சில கூட்டுத் திட்டங்கள், குறிப்பாக விக்சனரி

மற்றும் காலின்ஸ் ஆன்லைன் அகராதி. மேலும் சில, நகர்ப்புற அகராதியைப் (Urban Dictionary) போலவே, பயனர்களால் வழங்கப்பட்ட பதிவுகள் (சில நேரங்களில் சுய-முரண்பாடானவை) கொண்டிருக்கும். சிறப்பு நோக்கங்களுக்காகப் பல அகராதிகள், குறிப்பாகத் தொழில்முறை மற்றும் வர்த்தக சொற்கள் மற்றும் பிராந்திய பேச்சுவழக்குகள் மற்றும் மொழி வேறுபாடுகள், நிறுவனங்கள் மற்றும் தனிப்பட்ட ஆசிரியர்களின் வலைத்தளங்களில் வெளியிடப்படுகின்றன. அவை பெரும்பாலும் தேடல் செயல்பாடு இல்லாமல் பட்டியல் வடிவத்தில் வழங்கப்படலாம் என்றாலும், தகவல்களைச் சேமித்து அனுப்பும் முறையின் காரணமாக, அவை மின்னணு அகராதிகள் ஆகும்.

மதிப்பீடு

வன்பொருள் (கையடக்க சாதனங்கள்), மென்பொருள் (விளக்கக்காட்சி மற்றும் செயல்திறன்) மற்றும் அகராதி உள்ளடக்கத்தின் தரத்தில் வேறுபாடுகள் உள்ளன. சில கையடக்கங்கள் மற்றவற்றை விட வலுவாகக் கட்டமைக்கப்பட்டுள்ளன, மேலும் விசைப்பலகைகள் அல்லது தொடுதிரை உள்ளீட்டு அமைப்புகளை வாங்குவதற்கு முன் உடல் ரீதியாக ஒப்பிடப்பட வேண்டும். கணினி அடிப்படையிலான அகராதி மென்பொருளின் GUI பற்றிய தகவல் சிக்கலான மற்றும் இரைச்சலானது முதல் எழுத்துரு அளவு மற்றும் வண்ணம் உள்ளிட்ட பயனர் வரையறுக்கக்கூடிய விருப்பங்களுடன் தெளிவான மற்றும் பயன்படுத்த எளிதானதாக இருக்கும்.

சொல்சார்/லெக்சிகல் தரவுத்தளத்தின் தரம் ஒரு முக்கியக் கருத்தாகும். கல்லூரி மற்றும் தொழில்முறை பயன்பாட்டிற்காக வடிவமைக்கப்பட்ட அகராதிகளில் பொதுவாகத் தரமான அச்சிடப்பட்ட அகராதியில் எதிர்பார்க்கப்படும் பெரும்பாலான அல்லது அனைத்து சொல்சார்/லெக்சிகல் தகவல்களும் அடங்கும். அச்சிடப்பட்ட அகராதிகளின் முன்னணி வெளியீட்டாளர்களுடன் இணைந்து உருவாக்கப்பட்ட மின்னணு அகராதிகளின் உள்ளடக்கம் பயணி அல்லது சாதாரண பயனரை இலக்காகக் கொண்டவை மிகவும் நம்பகமானவை, அதே சமயம் ஒவ்வொரு மொழிக்கும் தாய்மொழி

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

அகராதியியலார்களின் குழுக்களால் எழுதப்படாத இருமொழி அகராதிகள் கல்வி வேலைக்குப் பொருத்தமானவை அல்ல. சில உருவாக்குபவர்கள்/டெவலப்பர்கள் தங்கள் தயாரிப்புகளை CALICO போன்ற ஒரு சுதந்திரமான கல்வி அமைப்பால் மதிப்பீடு செய்ய விரும்புகின்றனர்.

மற்றுமொரு முக்கிய கருத்து என்னவென்றால், சாதனங்களும் அவற்றிலுள்ள அகராதிகளும் பொதுவாக ஒரு குறிப்பிட்ட சந்தைக்காக வடிவமைக்கப்பட்டுள்ளன. எடுத்துக்காட்டாக, கிட்டத்தட்ட அனைத்து கையடக்க ஜப்பானிய-ஆங்கில மின்னணு அகராதிகளும் ஆங்கிலம் கற்கும் மற்றும் பயன்படுத்தும் ஜப்பானிய மொழியில் சரளமாக பேசுபவர்களுக்காக வடிவமைக்கப்பட்டுள்ளன; எனவே, ஜப்பானிய சொற்கள் பொதுவாக ஃபிரிகானா உச்சரிப்பு அர்த்தங்களை உள்ளடக்குவதில்லை, ஏனெனில் வாசகர் ஜப்பானிய மொழியில் கல்வியறிவு பெற்றவர் என்று கருதப்படுகிறது (எனினும் பதிவுகளின் தலைப்புச் சொற்களுக்கு உச்சரிப்பு உள்ளது). மேலும், சொற்களைத் தேடுவதற்கான முதன்மையான முறை உச்சரிப்பு ஆகும், இது தெரியாத உச்சரிப்புடன் ஒரு சொல்லைப் பார்ப்பதை கடினமாக்குகிறது (எடுத்துக்கட்டா, 網羅 "விரிவான" என்பது 毛うら, mōura என்று உச்சரிக்கப்படுகிறது என்பதைத் தெரிந்து கொள்ள வேண்டும்). இருப்பினும், ஜப்பானிய எலக்ட்ரானிக் அகராதிகளில் (முதன்மையாக சமீபத்திய மாடல்களில்) எழுத்து அங்கீகாரம் உள்ளது, எனவே பயனர்கள் (ஜப்பானிய மொழி பேசுபவர்களோ இல்லையோ) காஞ்சியை எழுதுவதன் மூலம் வார்த்தைகளைத் தேடலாம்.

இதே போன்ற வரம்புகள் பெரும்பாலான இரண்டு அல்லது பல மொழி அகராதிகளில் உள்ளன, மேலும் மொழிகள் ஒரே ஸ்கிரிப்ட் அல்லது எழுத்துக்களில் எழுதப்படாதபோது குறிப்பாக முடங்கும். பயனரின் தாய்மொழிக்கு உகந்த ஒரு அகராதியைக் கண்டறிவது முக்கியம்.

ஒருங்கிணைந்த தொழில்நுட்பம் (Integrated technology)

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

மின்னணு அகராதி மென்பொருளை இயக்கும் முறைமைகளை/கணினிகளை உருவாக்குபவர்கள்/டெவலப்பர்கள் பலர் பல்வேறு மொழி அடிப்படையிலான (அகராதி, மொழிபெயர்ப்பு, வரையறைவிளக்கங்கள், ஒருபொருள்பன்மொழிகள் மற்றும் எழுத்துப்பிழை சரிபார்ப்பு மற்றும் இலக்கண திருத்தம்) செயல்பாடுகளில் இருந்து நிரல்கள் வரை மற்றும் Google பயன்படுத்தும் AJAX API போன்ற வலைச் சேவைகளைச் சேர்க்க API மற்றும் SDK என்ற மென்பொருள் மேம்பாட்டு கருவிப்பெட்டிகளை (Development Kit tools) வழங்குகிறார்கள். இந்தப் பயன்பாடுகள் பல்வேறு வழிகளில் மொழியைக் கையாளுகின்றன, அகராதி/மொழிபெயர்ப்பு அம்சங்களையும், பொருண்மையியல்சார் தேடலுக்கான அதிநவீன தீர்வுகளையும் வழங்குகிறது. அவை பெரும்பாலும் சி++ ஏபிஐ (C++ API), எக்ஸ்எம்எல்-ஆர்பிசி சர்வர் (XML-RPC server), நெட் ஏபிஐ (NET API) அல்லது பைதான் ஏபிஐ (Python API) என பல இயக்க முறைமைகள் (மேக், விண்டோஸ், லினக்ஸ் போன்றவை) மற்றும் உருவாக்க/டெவலப்மெண்ட் சூழல்களில் கிடைக்கின்றன, மேலும் பிற வகையான தரவு அட்டவணைப்படுத்தலுக்கும் பயன்படுத்தப்படலாம்.

ஒருசில ஆன்லைன் தமிழ் அகராதிகள்

ஒரு காலத்தில் தமிழ் (தமிழ்-ஆங்கிலம்) அகராதி என்றால் சென்னை பல்கலைக்கழகத்தின் அகராதி, லிப்கோ தமிழ்-ஆங்கில அகராதி போன்றவை தான். இணையத்தில் எதுவுமே இருக்காது. இருபத்தைந்து வருடங்களாகவே ஆங்கில அகராதிகள் கணினியில் சுலபமாக கிடைக்கிறது. சுமார் பத்தாண்டுகளாக இணையத்தில் கூகுள், அமேசான் கிண்டல் போன்றவர்கள் தருகிறார்கள். ஆனால் தமிழ் அகராதிகள், அதுவும், பிழையில்லாத, நிரம்ப வார்த்தைகள் இருக்கும் அகராதிகள் இணையத்தில் இருக்கவில்லை. பல முயற்சிகள் நடந்தும், நிலைக்கவில்லை.

ஆனால் கடந்த சில ஆண்டுகளாகப் பல நல்ல தமிழ் அகராதிகள் இணையத்தில் சுலபமாகக் கிடைக்கிறன. அமேசான் கிண்டலில் புத்தகம் வாசிக்கும் போது ஒரே ஒரு தொடுதலில் மந்திரம் போல், தமிழ் அர்த்தம் தோன்றுகிறது. மைக்ரோசாஃப்ட் வேர்ட்டில் (Microsoft Word) கூட இரண்டு சொடுக்கில் வருகிறது. கூகிள் குரோம் (Google Chrome) அல்லது மைக்ரோசாஃப்ட் எட்ஜ் (Microsoft Edge) உலாவிகளில் நிறுவக்கூடிய கூகிள் அகராதியின் (Google Dictionary extension) உதவியால் ஒரு சுட்டியில் அர்த்தங்கள் தெரிகிறது. இப்போதெல்லாம் ஆங்கில வார்த்தையை (எ.கா.. “Smartphone in Tamil”) என்று கூகுளில் தேடினாலே தமிழ் வார்த்தை வந்துவிடுகிறது. ஆனால் பல சமயங்களில் கூகுள் செய்யும் பரிந்துரை நமக்குச் சரியாகப்படாது. இதற்காகவே நாம் சில அகராதிக் தளங்களைத் தேடவேண்டி வருகிறது..

இந்த அமெரிக்க பெரிய நிறுவனங்களைத் தவிர உள்ளூரிலும் பலர் (ஆத்ம நிர்பார் பாரத் / சுய சார்பு இந்தியா) பல தமிழ் தன்னார்வலர்கள், மற்றும் தமிழக அரசும் பலவற்றைப் படைத்து இலவசமாக வெளியிட்டுள்ளார்கள் - அதிலும் சில, படைப்பாக்கப் பொதுமங்கள் (Creative Commons) உரிமத்தில் கிடைக்கிறது - அவர்கள் அனைவருக்கும் நம் பாராட்டுக்கள். இந்தத் துறையை பல ஆண்டுகளாக நாம் பார்த்து வருகிறோம் என்கிற முறையில் நமக்கு மிக மகிழ்ச்சி. இதை இப்படித் தான் திறந்தவெளியில், விலையில்லாமல் செய்ய முடியும்.

நமக்குத் தெரிந்து இணையத்தில் இருக்கும் சில தமிழ் மற்றும் தமிழ்-ஆங்கிலம் அகரமுதலிகள் சில பட்டியலிடப்பட்டு சில விளக்கங்களும் தரப்பட்டுள்ளன.

- **தமிழக அரசின் சொற்குவை:** நாம் தினம் அடிக்கடி பயன்படுத்துவது, சில ஆண்டுகளுக்கு முன் தமிழக அரசின் தமிழ் வளர்ச்சித்துறை சார்பாக நிறுவப்பட்ட சொற்குவை தளம். இங்கே ஒரே தேடுதலில், ஆங்கில வார்த்தைக்கோ, தமிழ் வார்த்தைக்கோ பல திறவுகளிலிருந்து விளக்கமும், ஒருபொருள்பன்களையும்

(synonyms) கொடுக்கிறார்கள். தளம் வேகமாகவும் இருக்கிறது. இதில் புதுப் புது வார்த்தைகளை அடிக்கடி சேர்க்கவும் செய்வதாக டிவிட்டரில் வருகிறது. பாராட்டுக்கள். அரசைக் குறைமட்டுமே செய்யும் நாம், அவர்களின் நல்ல முயற்சிகளையும் பாராட்ட வேண்டும்.

- தமிழ் இணையக் கல்விக்கழகம் – இதில் சென்னைப் பல்கலைக்கழகத் தமிழ்ப் பேரகராதி, பால்ஸ் அகராதி மற்றும் பல அகராதிகள் இருக்கிறது
- விக்கிமூலம் தமிழ் அகராதிகளைத் தொகுப்பு
- அகரமுதலி.COM: உள்ளே எச். டி.எம். எல். லில் தெரியும் இணைய முகவரியைப் பார்த்துப் பயன்பட வேண்டாம், இது சரியானது தான். இருப்பது தமிழில், நீங்கள் இதை அகரமுதலி.COM என உள்ளீடுச் செய்தால் அதுவாகவே மாறிவிடும்
- மதன் கார்க்கியின் சொல் – இதைப் பயன்படுத்த இலவசப் பதிவு தேவை.
- தமிழ்ப் பேழை
- வாணி சொல்திருத்தியின் அகராதி
- அகராதி.com
- European dictionary for world languages

தமிழ் இணையக் கல்விக்கழக அகராதிகள்:

இந்தப் பிரிவின் கீழ் தமிழ்-தமிழ், தமிழ்-ஆங்கிலம், தமிழ் நிகண்டு போன்ற பல அகராதிகள் இடம்பெறுகின்றன. கீழ் காணும் அகராதிகள் இணைய தளத்தில் இடம் பெற்றுள்ளன:

- சென்னைப் பல்கலைக்கழகத் தமிழ்ப் பேரகராதி
- பால்ஸ் அகராதி
- சென்னைப் பல்கலைக்கழக ஆங்கில-தமிழ் அகராதி
- தமிழ்-தமிழ் அகர முதலி
- தமிழ் இணையக் கல்விக்கழகத்தின் உச்சரிப்புடன் கூடிய மின் அகராதி
- சங்க இலக்கிய அகராதி (பாட்டும் தொகையும்)

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

- காலக் குறிப்பு அகராதி
- செந்தமிழ் சொற்பிறப்பியல் பேரகரகமுதலி

மேலும் சில:

- இராம கி ஐயாவின் வளவு
- திரு மணவை முஸ்தபா அவர்களின் கணினி களஞ்சிய அகராதி, இதன் இரண்டு பாகமும் தமிழ் இணையக் கல்விக்கழகத்தின் தளத்தில் விலையில்லாமல் கிடைக்கிறது: பாகம் ஒன்று, பாகம் இரண்டு
- How to get Tamil meaning in Microsoft Word?
- How to get Tamil meaning in Google Search?
- Google Dictionary Chrome Extension

இணையதளத்தில் புதிய இலவச தமிழ் அகராதி: அண்ணா பல்கலைக. ஆசிரியர்களின் உருவாக்கம்

தமிழ் மக்களின் பயன்பாட்டிற்காக இணையதளத்தில் இலவச தமிழ் அகராதியை உருவாக்கியுள்ளது அண்ணா பல்கலைக்கழகம். அண்ணா பல்கலைக்கழக ஆசிரியரும், கவிஞர் வைரமுத்துவின் மகனுமான மதன் கார்க்கி, ஆசிரியர்கள் கீதா, ஷோபா, ரஞ்சனி ஆகியோரும் சேர்ந்து இணையதளத்தில் இலவச தமிழ் அகராதியை உருவாக்கியுள்ளனர். இதன் முகவரி www.agaraadhi.com

ஏற்கனவே, இணையதளத்தில் பல தமிழ் அகராதிகள் இருக்கின்றன. ஆனால், அவை அளித்திடாத 20 சேவைகளை இந்தப் புதிய அகராதி அளிக்கிறது. தமிழ்ச் சொல் உருவாக்கம், பிழைத்திருத்தம், மாற்றுச் சொற்கள், சொற்கள் பயன்பாடு, திருக்குறள் பயன்பாடு, பாரதியார் மற்றும் ஒளவையார் பாடல்களில் உள்ள சொற்களின் பயன்பாடு உள்ளிட்ட பல பயன்பாடுகள் இதில் உள்ளன. குழந்தைகள் முதல் பெரியவர்கள் வரை அனைவரும் எளிதாகப் பயன்படுத்தும் வகையில் இந்த அகராதி அமைக்கப்பட்டுள்ளதாகத் தெரிவிக்கப்பட்டுள்ளது.

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

சில சொற்களுக்குப் பொருத்தமான தமிழ்த் திரைப்படப் பாடல்களின் வரிகளையும் இதனுடன் இணைத்துள்ளது வித்தியாசமாக உள்ளது.

அதிகம் படிக்க <https://tamil.oneindia.com/art-culture/essays/2010/18-tamil-online-dictionary-anna-university.html>

ஆன்லைனில் பல தமிழ் அகராதிகள் குறிப்பாக ஆங்கிலத்திலுள்ள சொற்களுக்குத் தமிழில் அர்த்தம் தருவனவாகக் கிடைக்கப்பெறுகின்றன. சிலவற்றைப் பதிவிறக்கம் செய்து ஆன்லைனில் பயன்படுத்திக்கொள்ளலாம். ஒரு சில அகராதிகள் பற்றி இங்கு சுருக்கமாகத் தரப்பட்டுள்ளன.

1.TAMIL DICTIONARY என்ற பெயரில் கிடைக்கபெறும் ஒரு ஆன்லைன் அகராதி பற்றி பின்வரும் குறிப்பு காணப்படுகின்றது:

தமிழ் அகராதி உலகின் சிறந்த ஆன்லைன் தமிழ் அகராதி. சொற்களின் விரிவான விளக்கத்துடன் 100,000+ வார்த்தைகள் சேகரிப்பு, இது உலகின் மிகப்பெரிய சேகரிப்புகளில் ஒன்றாகும். மேலும் எங்கள் அகராதியில் தொடர்ந்து புதிய வார்த்தைகளைச் சேர்க்கிறோம்.

அகராதியில், மெட்ராஸ் பல்கலைக்கழக அகராதி, ஜே.பி. ஃபேப்ரிசியஸ் தமிழ் மற்றும் ஆங்கில அகராதி, மிரான் வின்ஸ்லோ - ஒரு விரிவான தமிழ் மற்றும் ஆங்கில அகராதி, டேவிட் மெகால்பின் - தமிழ் மற்றும் தமிழ் அகர முதுலிகளுக்கான முக்கிய சொற்றொகை உட்பட பல அகராதி ஆதாரங்கள் உள்ளன.

முதல் முறையாக நீங்கள் ஒரு தமிழ் அகராதிக்கு REST APIஐப் பயன்படுத்த எளிதானது. உங்கள் சொந்தப் பயன்பாடுகளை உருவாக்க எங்கள் தரவைப் பயன்படுத்த விரும்பினால், எங்கள் தமிழ் அகராதி APIஐப் பார்க்கவும்.

2.Tamilcube.com - Best online translation and educational resources

மேற்கண்ட பெயரில் அமையும் ஆன்லைன் அகராதி பின்வரும் வசதிகள் கொண்டுள்ளதாக அதன் முகப்பில் தரப்பட்டுள்ளது;

எழுத்துப்பிழை சரிபார்ப்புடன் உலகின் மிகப்பெரிய மற்றும் மிகவும் பிரபலமான இலவச ஆங்கிலத்திலிருந்து தமிழ் அகராதி & தமிழிலிருந்து ஆங்கிலம் அகராதிக்கு வரவேற்கிறோம்! தமிழில் இருந்து ஆங்கில மொழி பெயர்ப்பு, ஆங்கிலத்தில் இருந்து தமிழ் மொழி பெயர்ப்பு அல்லது எண்கள் தமிழ் வார்த்தை மாற்றத்தை விரைவாகத் தேட இந்த அகராதி உதவுகிறது. இந்த அகராதியில் தேடலுக்கான வார்த்தைகளின் எண்ணிக்கை, ஏற்கனவே 500,000ஐ எட்டியுள்ளது, இன்னும் வளர்ந்து வருகிறது. நீங்கள் இதை ஒரு சொற்களஞ்சியமாகவும் (Thesaurus) பயன்படுத்தலாம். உங்கள் தேடலுக்கான பொருத்தமான தமிழ் சொற்களை வழங்குவதோடு, இந்த அகராதி தொடர்புடைய தமிழ் சொற்களையும் அவற்றின் உச்சரிப்புடன் வழங்குகிறது.

3 முறைகளில் அகராதி : இந்த அகராதியை நீங்கள் மூன்று வழிகளில் பயன்படுத்தலாம்: ஆங்கில சொற்களைத் தமிழில் மொழிபெயர்க்கவும், தமிழ் சொற்களை ஆங்கிலத்தில் மொழிபெயர்க்கவும், எண்களைத் தமிழ் சொற்களுக்கு மொழிபெயர்க்கவும். இந்த தளத்தின் நோக்கம் தமிழ்ச் சொற்கள் மற்றும் தமிழ் எண்களை எளிதாகக் கற்க உதவுவதாகும்.

iGoogle : இப்போது, இந்த அகராதி கேஜெட்டை உங்கள் iGoogle இல் சேர்க்கலாம்!

மொபைல் தமிழ் அகராதி

உங்கள் ஃபோன்கள் மற்றும் ஆண்ட்ராய்டு ஃபோன்கள், iPhone, iPad, iPod Touch அல்லது Blackberry உள்ளிட்ட மொபைல் சாதனங்களில் இயங்குவதற்கு இந்த அகராதி மிகவும் பொருத்தமானது. உங்கள் கைப்பேசியிலிருந்து இந்த தமிழ் அகராதி வலைப்பக்கத்திற்குச் சென்று தேடத் தொடங்குங்கள். கூகுள் ஆண்ட்ராய்டு ஆப்ஸ் அல்லது ஆப்பிள் ஐபோன் ஆப்ஸ் எதையும் பதிவிறக்கம் செய்ய வேண்டிய அவசியம் இல்லை. ஆயிரக்கணக்கான பார்வையாளர்கள் இந்த தமிழ் அகராதியைத் தங்கள் ஆண்ட்ராய்ட் ஸ்மார்ட் மொபைல் போன்கள், ஐபோன், ஐபாட், ஐபாட் டச் அல்லது பிளாக்பெர்ரி ஆகியவற்றிலிருந்து தினமும் தேடுகிறார்கள்!

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

ஆங்கிலத்திலிருந்து தமிழ் மொழி பெயர்ப்பு அகராதி

ஆங்கிலத்திலிருந்து தமிழ் மொழிபெயர்ப்பிற்கு, மேலே உள்ள தேடல் பெட்டியில் நீங்கள் தமிழில் மொழிபெயர்க்க விரும்பும் ஆங்கில வார்த்தையை உள்ளிட்டு 'SEARCH' என்பதைக் கிளிக் செய்யவும்.

தமிழ் ஆங்கில மொழிபெயர்ப்பு அகராதி

தமிழில் இருந்து ஆங்கில மொழிபெயர்ப்பிற்கு, மேலே உள்ள தேடல் பெட்டியில் தமிழ் வார்த்தைகளை உள்ளிட பல விருப்பங்கள் உள்ளன.

1. மேலே உள்ள பெட்டியில் உங்கள் தமிழ் வார்த்தைகளை (யூனிகோடில்) வெட்டி ஒட்டவும் மற்றும் 'தேடு' என்பதைக் கிளிக் செய்யவும். யூனிகோட் தமிழில் தட்டச்சு செய்ய எங்கள் தமிழ் மொழிபெயர்ப்பாளரை நீங்கள் பயன்படுத்தலாம்.
2. உங்களுக்கு Romanised Transliteration தெரிந்திருந்தால், மேலே உள்ள Translate Unicode Tamil to English என்ற பொத்தானைத் தேர்ந்தெடுத்து ஆங்கிலத்தில் தட்டச்சு செய்யத் தொடங்கலாம். நீங்கள் ஆங்கில எழுத்துக்களை ஒலிப்பு முறையில் தட்டச்சு செய்து, ஸ்பேஸ் பாரை அழுத்தினால், இவை தானாகவே தமிழ் எழுத்துக்களாக மாற்றப்படும். எ.கா., ammaa என்று ஆங்கிலத்தில் டைப் செய்து ஸ்பேஸ் பாரை அழுத்தினால் அது அம்மா என மாறும்.

எண்களை தமிழ் வார்த்தைகளாக மாற்றவும்

எண்களைத் தமிழ் சொற்களாக மாற்ற, Translate Number to Tamil Word பொத்தானைத் தேர்ந்தெடுத்து, மேலே உள்ள தேடல் பெட்டியில் எண்ணை உள்ளிட்டு 'SEARCH' என்பதைக் கிளிக் செய்யவும். காற்புள்ளி போன்ற பிரிப்பான்களைப் பயன்படுத்த வேண்டாம். எடுத்துக்காட்டாக, நீங்கள் 555 ஐ அழுத்தி, SEARCH ஐக் கிளிக் செய்தால், இது ஐநூற்று ஐம்பத்து ஐந்து என மொழிபெயர்க்கப்படும். அனுமதிக்கப்பட்ட அதிகபட்ச எண்ணிக்கை 999999999 (கிட்டத்தட்ட ஒரு பில்லியன்). எங்கள்

அகராதியின் இந்த அம்சம் தமிழ் எண்களை மிக விரைவாகக் கற்றுக்கொள்ள உதவுகிறது.

3. SHABDKOSH -English and Indian Language Dictionaries

ஆங்கில தமிழ் அகராதி மற்றும் மொழிபெயர்ப்பு

இந்தத் தளம் ஆங்கிலத்திலிருந்து தமிழ் அகராதியையும், தமிழிலிருந்து ஆங்கில அகராதியையும் வழங்குகிறது. 2003ஆம் ஆண்டு தொடங்கப்பட்ட இந்த தளம் தற்போது உலகம் முழுவதும் நூற்றுக்கும் மேற்பட்ட நாடுகளில் மில்லியன் கணக்கான மக்களால் பயன்படுத்தப்படுகிறது.

அகராதிகளும் தொழில்நுட்பமும்

1.அகராதியியலில் தரவுதொகுதிப் புரட்சி (The Corpus Revolution in Lexicography)

மொழியை விவரிப்பதில் உண்மையான மொழி தரவுகளுடன் பணிபுரிவதன் முக்கியத்துவத்தை அகராதியியலாளர்கள் எப்போதும் புரிந்துகொண்டுள்ளனர். கணினிகளின் வருகைக்கு முன், தீவிர அகராதி உருவாக்கம் என்பது இலக்கியத்திலிருந்து மில்லியன் கணக்கான மேற்கோள்களை கைமுறையாக சேகரிக்கும் கடினமான செயல்முறையை உள்ளடக்கியது. சிறப்பு வாசிப்பு நிகழ்ச்சிகள் மூலம் படித்த மக்களால் சில நேரங்களில் அகராதி தயாரிப்பாளர்கள் இந்த பணியில் உதவினார்கள். இதன் விளைவாக மேற்கோள்கள் மேற்கோள் சீட்டுகளில் வைக்கப்பட்டு, பெரிய கோப்புகளில் சிரமமின்றி வரிசைப்படுத்தப்பட்டன. இந்த முறை மிகவும் கடினமானதாக இருந்தது. இதன் விளைவாக, கையேடு மேற்கோள்களின் எந்தவொரு தரவுத்தளமும் மொழியின் ஆக்கப்பூர்வமான பயன்பாட்டின் நிகழ்வுகளை வலியுறுத்த முனைகிறது, ஆனால் பொதுவான சொற்களின் ஆர்வமற்ற அன்றாட பயன்பாடுகள் குறிப்பிடப்படாமல் உள்ளன, ஏனெனில் அவை பதிவு செய்யத் தகுதியற்றவையாகத் தோன்றுகின்றன.

1980-களில், முன்னோடியான COBUILD திட்டத்திற்கு (Sinclair, 1987) நன்றி சொல்லும் வகையில் அகராதி உருவாக்கம் ஒரு பெரிய புரட்சியைச் சந்தித்தது.

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

பனுவல்/உரைத் தரவுத்தொகுக்களை (text corpora) முறையாகப் பயன்படுத்துவதற்கான முதல் அகராதி திட்டம் இதுவாகும், மேலும் தரவுத்தொகுதிப் புரட்சி கற்பவர்களின் அகராதிகளுடன் (learners' dictionaries) தொடங்கப்பட்டது. அங்கிருந்து, அது படிப்படியாக பிற வகை அகராதிகளுக்கு பரவியது மற்றும் தரவுத்தொகுதி மொழியியலின் (corpus linguistics) வளர்ச்சியைத் தொடங்கியது.

COBUILD குழு, அகராதியின் தொகுப்பிற்காக 7.3 மில்லியன் சொற்களின் மின்னியல் தொகுப்பை சேகரித்தது, மேலும் இந்த எண்ணிக்கை மேலும் உரையைச் சேர்த்து, இறுதி எட்டிங்/திருத்தக் கட்டத்தில் 18 மில்லியன் சொற்களை எட்டியது. தரவுத்தொகுதி - ஆரம்பத்தில் பர்மிங்காம் ஆங்கில உரையின் சேகரிப்பு (Birmingham Collection of English Text) என்று அறியப்பட்டது, பின்னர் ஆங்கிலேய வங்கி (Bank of English) என்று மறுபெயரிடப்பட்டது - அந்த நேரத்தில் மிகப்பெரியதாகத் தோன்றியது, ஆனால் பில்லியன் கணக்கான சொற்களை வைத்திருக்கும் இன்றைய தரவுத்தொகுதிகளுடன் (corpora) ஒப்பிடும்போது, இது மிகவும் சிறியது.

அகராதி உருவாக்கத்தில், தரவுத்தொகுதிகள் பல வழிகளில் பயனுள்ளதாக இருக்கும். அவை சாத்தியமான தலைப்புச் சொற்களைத் (headwords) தேர்ந்தெடுப்பதற்கும், அர்த்தங்கள் மற்றும் பயன்பாடுகளை அடையாளம் காண்பதற்கும் அடிப்படையாக அமைகின்றன. அவை சொற்களின் உருபனியல் மற்றும் தொடரியல் நடத்தை மற்றும் மாற்று எழுத்து வடிவங்களின் (alternative spelling forms) ஒப்பீட்டு அதிர்வெண் ஆகியவற்றின் விளக்கத்திற்கான புறநிலை தரவை வழங்குகின்றன. சேர்ந்துவருகை/சொல்லடிவகைப்பாட்டு நடத்தை (collocational behavior) மற்றும் குறிப்பிடத்தக்க பல்-சொல் அலகுகளை (multi-word units) அடையாளம் காண்பது மிகவும் மேம்பட்ட பயன்பாடுகளில் அடங்கும் இதற்குப் பெரிய

அளவிலான தரவுத்தொகுதிகள் மற்றும் அதிக நுட்பமான மொழிப் பொறியியல் கருவிகள் தேவைப்படுகின்றன. பனுவல் தரவுத்தொகுதிகள், அகராதியலார்களுக்கு ஏராளமான மொழியின் உண்மையான பயன்பாட்டின் சாத்தியமான எடுத்துக்காட்டுகளுக்குத் தயாராக அணுகலை வழங்குகிறது, உண்மையில் COBUILD-இன் அசல் விற்பனைப் புள்ளி அது உண்மையான மொழியைக் கையாள்வதாகும். எவ்வாறாயினும், ஒரு அகராதி உள்ளீட்டில் சேர்க்க ஒரு எடுத்துக்காட்டு பொருத்தமானது என்பதற்கு நம்பகத்தன்மை மட்டுமே உத்தரவாதம் இல்லை என்பதை தரவுத்தொகுதி அகராதியியலாளர்கள் (corpus lexicographers) இன்று உணர்ந்துள்ளனர்.

COBUILD அதன் வழிமுறையைத் தரவுத்தொகுதி-உந்துதல் (corpus-driven) என்று விவரிக்கிறது. இந்தச் சொல் முக்கியமாக விதிவரு அணுகுமுறையைக் (inductive approach) குறிக்கிறது, அங்கு ஒருவர் ஆதாரத்துடன் தொடங்குகிறார். இந்த அணுகுமுறை சில நேரங்களில் தரவுத்தொகுதி-அடிப்படையாக (corpus-based) பண்பாக்கம் செய்யப்படும் ஒன்றுக்கு எதிரானது, இதில் தரவுத்தொகுதி குறைவான மைய, கூடுதல் நிரப்புப் பங்கை வகிக்கிறது, பெரும்பாலும் முன்பே இருக்கும் யோசனைகளுக்கான (பிந்தைய தற்காலிக) ஆதாரமாக உள்ளது. அகராதியியல் துறையில், ஒரு திட்டம் ஒரு தரவுத்தொகுதியை ஓரளவு நம்பியிருக்கலாம், ஆனால் பிற சுதந்திரமான கருதல்களையும் உள்ளடக்கியிருக்கலாம். எடுத்துக்காட்டாக, ஒரு பதிவுக்குள் அர்த்தங்களை வரிசைப்படுத்துவதில், அகராதியியலாளர்கள் குறிப்பிட்ட அடையாளம் காணக்கூடிய பயன்பாடுகளின் புறநிலை நிகழ்வெண்ணால் வழிநடத்தப்படாமல், எந்த அர்த்தம் பொருண்மையியல் அடிப்படையில் மிகவும் அடிப்படையானது என்பதைக் கருத்தில் கொள்ளலாம். எனவே, summit என்பதற்கு 'மலையின் உச்சி' என்ற அர்த்தம் முதலில் பட்டியலிடப்படலாம், 'முக்கியமான அரசியல்

சந்திப்பு' என்ற அர்த்தம் முந்தையதில் இருந்து பெறப்பட்டது என்ற நிலையில் பின்னர் பட்டியலிடப்படலாம். ஆனால் COBUILD-இல், கார்பஸ் கொள்கையின் முதன்மையால் கட்டளையிடப்பட்டபடி, உரையில் அடிக்கடி நிகழும் அரசியல் அர்த்தம் முதலில் பட்டியலிடப்பட்டுள்ளது.

மொழி கற்பவர்களுக்கான அகராதிகளை உருவாக்கும் அகராதியியலாளர்களும் கற்கும் தரவுத்தொகுதியைப் பயன்படுத்தலாம். இவை பல்வேறு நிலைகளில் உள்ள மொழி கற்பவர்களால் எழுதப்பட்ட (அல்லது, பொதுவாக, குறைவாகப் பேசப்படும்) தாய்மொழி அல்லாத உரைகளின் தொகுப்புகளாகும். ஒரு மொழி அனுபவத்தைக் கற்கும் குறிப்பிட்ட சிக்கல்களைக் கண்டறிவதற்காக, ஒரு கற்றல் தரவுத்தொகுதியை அகராதியியல் அடிப்படையில் ஆராயப்படலாம், இதனால் பொருத்தமான பதிவுகளில் உள்ள சிக்கலான இடங்களுக்கு, பொருத்தமான எடுத்துக்காட்டுகளைத் தேர்ந்தெடுப்பதன் மூலம் அல்லது ஒரு பயன்பாட்டில் உள்ள சிக்கலைப் பயன்பாட்டுப் பெட்டியில் வெளிப்படையாகக் கூறுவதன் மூலம் கவனத்தை ஈர்க்க முடியும்.

2. தரவுத்தொகுதி வினவல் அமைப்புகள்

தரவுத்தொகுதி, பயனுள்ளதாக இருக்க, தரவுத்தொகுதி பயனர்கள் உரை சேகரிப்பை எளிதாக விசாரிக்கக்கூடிய முன்-இறுதி இடைமுகத்துடன் பொருத்தப்பட்டிருக்க வேண்டும். தரவுத்தொகுதித் தரவை வழங்குவதற்கான தரமான வழி, ஒரு உரைச் சூழலில் இலக்குச் சொல்லைக் (முக்கியச்சொல்) காட்டும் தொடரடைவு வரிகள் (concordance lines) வழியாகும், பொதுவாக ஒரு உரையின் ஒற்றை வரி. அசல் COBUILD திட்டத்தில், தனிப்பட்ட சொற்களுக்கான தொடரடைவுகள் காகிதத்தில் அச்சிடப்பட்டன, ஏனெனில் கணினிகள் நிகழ்நேரத்தில் தொடரடைவுகளை உருவாக்க மிகவும் தகுதியற்றதாக இருந்தன.

உரைத் தரவுத்தொகுதிக்கான (text corpora) இன்றைய இடைமுகங்கள், பயன்படுத்தக்கூடிய பதிவை முடிந்தவரை திறமையாகவும், குறைந்தபட்ச முயற்சியுடனும் பெறுவதற்கு அகராதியியார்களுக்கு உதவுவதற்கான அதிநவீன வழிகளை வழங்குகின்றன. மிகவும் புதுமையானது சொல் விவரக்குறிப்பு மென்பொருளாகும் (word-profiling software), இது தேடல் சொற்களின் கட்டமைக்கப்பட்ட காட்சிகளை உருவாக்க வடிவமைக்கப்பட்டுள்ளது. ஆங்கிலத்திற்கான இந்த வகை இலவச ஆதாரம் ஜஸ்ட் த சொற் சேவை/Just the Word service (<http://www.just-the-word.com>) ஆகும். SketchEngine (Kilgariff & Tugwell, 2002) என்ற ஒரு முன்னணி வணிக அமைப்பானது இன்னும் அதிக நெகிழ்வுத்தன்மையைக் கொண்டுள்ளது; இது பல்வேறு வகையான விளக்கக்காட்சிகளை அனுமதிக்கிறது: தொடரடைவுகள் (concordances), தொடரடிவகைப்பாடுகள் (collocates), ஒருபொருள்பன்மொழிகள் (synonyms), ஒருபொருள்பன்மொழிய ஒப்பீடுகள், மேலும் இது வளர்ந்து வரும் மொழிகளுக்குக் கிடைக்கிறது (எ.கா. Radziszewski, Kilgariff, & Lew, 2011). இந்த ஒழுங்குமுறை உள்ளமைக்கப்பட்ட தரவுத்தொகுதியுடன் பொருத்தப்பட்டுள்ளது; கூடுதலாக, பயனர்கள் தங்கள் சொந்த தரவுத்தொகுதியை உருவாக்க முடியும். SketchEngine இன் மிகவும் பயனுள்ள அம்சம் சொல் சுருக்கப்படங்கள் (word sketches): ஒரு சொல்லின் இலக்கண மற்றும் சொல்லடிவகைப்பாட்டு நடத்தையின் (collocational behavior) ஒரு பக்க சுருக்கங்கள் .

பெருகிவரும் தரவுத்தொகுதியின் ஒரு விளைவு என்னவென்றால், உரைச் சான்றுகள் அதிகமாகி, விரிவாக ஆராய இயலாது. சிக்கலைத் தீர்க்க, தரவுத்தொகுதியில் உள்ள பல சாத்தியமான வாக்கியங்களிலிருந்து சிறந்த எடுத்துக்காட்டு வாக்கியங்களைப் பிரித்தெடுக்க மொழி தொழில்நுட்பம்

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

பயன்படுத்தப்படுகிறது. இதைச் செய்யும் ஒரு அமைப்பு GDEX (Kilgarriff, Husak, McAdam, Rundell, & Rychlý, 2008). அதன் வெளியீட்டின் (output) தரத்தை <http://forbetterenglish.com/>-இல் சோதிக்கலாம். இது போன்ற முயற்சிகள், ஒரு மனித அகராதியியலாருக்கு இயன்ற அளவு துர்நாற்றத்திலிருந்து விடுபடுவதை நோக்கமாகக் கொண்டுள்ளன, இதனால் அதிகபட்ச பணிகள் தானாகவே செய்யப்படுகின்றன (Kilgarriff, Kovář, & Rychlý, 2010).

3. இருமொழி அகராதியியலில் தரவுத்தொகுதிகள்

இருமொழி அகராதி திட்டத்தில், இருமொழி அகராதியைத் தொகுப்பது போன்று இரண்டு ஒற்றை மொழித் தரவுத்தொகுதிகள் பயன்படுத்தப்படலாம். மற்றொரு அணுகுமுறை சம்பந்தப்பட்ட இரண்டு மொழிகளின் ஒப்பிடக்கூடிய அல்லது இணையான தரவுத்தொகுதிகளைச் சேகரிப்பதை உள்ளடக்கியது. ஒப்பிடக்கூடிய தரவுத்தொகுதிகள் இரண்டு மொழிகளில் உள்ள உரைகளைக் கொண்டுள்ளது, அவை செயல்பாட்டு ரீதியாகச் சமமானவை மற்றும் ஒரே விகிதத்தில் உள்ளவை. இணைத் தரவுத்தொகுதிகள் உரை இணைகளைக்/ஜோடிகளைக் கொண்டுள்ளது, ஒன்று (குறைந்தது) பொதுவாக மொழிபெயர்ப்பாகும். இணைத் தரவுத்தொகுதிகள் வரிசைப்படுத்தப்படலாம் (ஒத்திசைக்கப்படலாம்), இதனால் அகராதியியலார் பொதுவாக ஒத்த சூழல்களில் என்ன சொற்கள் மற்றும் கட்டமைப்புகள் பயன்படுத்தப்படுகின்றன என்பதை விரைவாக மதிப்பிட முடியும். சிறப்பு இயற்கைமொழி ஆய்வுக் கருவிகளின் (NLP tools) உதவியுடன், மொழிகளுக்கு இடையே வழக்கமான சொல்லுக்குச் சொல் நிகரன்களைப் பிரித்தெடுக்க முயற்சி செய்யலாம். இது (அறிவியல்) சொற்களுக்குச் சிறப்பாகச் செயல்படும், மேலும் சிறப்பு அகராதிகளைத் தொகுப்பதில் பயனுள்ளதாக இருக்கும்.

4. அகராதி எழுதும் அமைப்புகள் (Dictionary Writing Systems)

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

உரை திருத்தி (text editor), தரவுத்தளப் பயன்பாடு (database application), வடிவமாற்றக் கருவிகள் (format conversion tools), பணிப்பாய்வு (workflow) மற்றும் திட்டமிடும் மென்பொருள்கள் (planning software) போன்ற தனித்த கருவிகளில் இருந்து அகராதி இயலாரின் பணிப்பெட்டியை (lexicographer's workbench) ஒன்றிணைப்பது முற்றிலும் சாத்தியம். இருப்பினும், பல எடிட்டர்களும் வெளியீட்டாளர்களும் இப்போது ஒரு அகராதி திட்டத்திற்கு தேவையான அனைத்து அல்லது பெரும்பாலான செயல்பாடுகளை ஒரு அகராதி எழுதும் அமைப்பு (Dictionary Writing System அல்லது DWS, சுருக்கமாக) என்ற ஒரு மென்பொருளில், ஒருங்கிணைப்பதில் இருந்து ஒரு நன்மையைப் பெறலாம் என்று நம்புகிறார்கள்.. வெளியீட்டாளர்கள் அல்லது குழுக்கள் தங்கள் சொந்த நோக்கங்களுக்கு ஏற்ப தனிப்பயனாக்கப்பட்ட DWS-ஐ உருவாக்கலாம் அல்லது ஏற்கனவே உள்ள ஒன்றை அவர்கள் மாற்றியமைக்கலாம். பல DWS தொகுப்புகள் இப்போது கிடைக்கின்றன, மிகவும் பிரபலமானவை: IDM DPS, TLex, iLEX, ABBY Lingvo Content (வணிக தொகுப்புகள்); மற்றும் Lexique Pro, Léacslann (வணிகமற்ற அமைப்புகள்). TLex/டி.லெக்ஸ் அகராதி எழுதும் மென்பொருளுக்கு ஒரு எடுத்துக்காட்டு. அகராதி எழுதும் அமைப்பு அகராதித் திட்டங்களை செயல்படுத்துவதில் உள்ள பல பணிகளுக்கு உதவும். இது பொதுவாக எக்ஸ்எம்எல்லில் (XML) பதிவுகளைத் (entries) திருத்த உதவுகிறது. அகராதிசார் உள்ளடக்கம் (Lexicographic content) (அல்லது தரவு) இப்போது பொதுவாக விளக்கக்காட்சியிலிருந்து (presentation) பிரிக்கப்படுகிறது, இது தனித்தனியாக கையாளப்படுகிறது, இதனால் ஒரு குறிப்பிட்ட தரவுப் புலம் (data field) எவ்வாறு காட்டப்படுகிறது என்பதைப் பாதிக்கும் முடிவுகள் எந்த நேரத்திலும் எடுக்கப்படலாம், மேலும் கொடுக்கப்பட்ட வகையின் அனைத்து தரவையும் தானாகவே பாதிக்கும்

(அர்த்தங்களுக்கு எண்ணிடுதல் மற்றும் அர்த்தங்களை உள்தள்ளுதல், சாய்வு எழுத்துக்களில் எடுத்துக்காட்டு வாக்கியங்களைக் காட்டுதல், தோட்டாக்களால் அவற்றைப் பிரித்தல் போன்றவை). ஒரு அதிநவீன முறைமை (state-of-the-art system) இறுதி விளக்கக்காட்சியை உண்மையான நேரத்தில் உருவகப்படுத்தும். பல-சொல் வெளிப்பாடுகளுக்கு, குறுக்குச் சரிபார்ப்புகளை வைத்திருக்க முடியும், இதனால் அகராதித் தரவு (lexicographic data) ஒருமுறை மட்டுமே உள்ளிடப்படும், ஆனால் ஒரே மாதிரியான பாணியில் தொடர்புடைய அனைத்து தலைப்புச்சொற்களின் (headwords) கீழும் காட்டப்படும். மைய சேவையகத்தில் (central server) எல்லாத் தரவையும் சேமிப்பதன் மூலமும், பயன்பாட்டில் உள்ள கோப்புகளைப் பூட்டுவதன் மூலமும் குழுப்பணி ஆதரிக்கப்படலாம். தலைப்புச் சொற்கள் பல்வேறு அளவுகோல்களைப் பயன்படுத்தி தொகுக்கப்படலாம் மற்றும் குறிப்பிட்ட நிபுணர்களுக்கு ஒதுக்கப்படலாம். இறுதியாக, புள்ளியியல் அறிக்கைகள் மூலம் பணிப்பாய்வு ஒழுங்கமைக்கப்பட்டு முன்னேற்றம் கண்காணிக்கப்படும்.

5. ஆன்லைன் அகராதிகளுக்கான புதிய பணிப்பாய்வு

தொழிநுட்பம் அகராதிகளைத் தொகுக்கும் வழிகளில் மட்டுமல்ல, அவை தொகுக்கப்பட்ட விதத்திலும் புரட்சியை ஏற்படுத்தியுள்ளது; இந்த நேரத்தில் பாரம்பரிய அச்சு ஊடகத்திலிருந்து மின்னணு ஊடகத்திற்கு மாறுவதை நாம் காண்கிறோம். உலகெங்கிலும் உள்ள அதிகமான மொழி கற்பவர்கள், கணினி அடிப்படையிலான பயன்பாடுகள், கையடக்க தனித்த தயாரிப்புகள் (குறிப்பாக ஆசியாவில் பிரபலமானது), ஆன்லைன் அகராதிகள், மொபைல்-ஃபோன் பயன்பாடுகள் அல்லது மின் புத்தக வாசகர்களில் அகராதிகள் போன்ற மின்னணு அகராதிகளைப் பயன்படுத்தத் தொடங்கியுள்ளனர். ஆன்லைன் அகராதிகள், குறிப்பாக, அகராதி தயாரிப்பதில் ஒரு புதிய அணுகுமுறைக்கு அழைப்பு விடுக்கிறது. அச்சிடப்பட்ட தலைப்புகளைப் போல

அகராதி தொகுப்பின் பாரம்பரிய நிலைகள் இனி பெற முடியாது. ஆன்லைன் அகராதிகளைத் தேவைப்படும் போது (தினசரி அடிப்படையில் கூட) அதிகரித்து மேம்படுத்த முடியும் என்பதால், வளர்ச்சி சுழற்சியில் அகராதி பயனர்களிடமிருந்து ஒரே நேரத்தில் கருத்துகள் இருக்கலாம் (De Schryver & Prinsloo, 2001). எடுத்துக்காட்டாக, பல பயனர்கள் ஒரு குறிப்பிட்ட பொருளைத் தேடத் தொடங்கினால், பாரம்பரிய புதிய பதிப்பிற்காகக் காத்திருக்காமல் உடனடியாக அகராதியில் சேர்க்கப்படலாம்.

6. அகராதி பயனர்களுக்குத் தொழில்நுட்பம் எவ்வாறு உதவுகிறது

நன்கு வடிவமைக்கப்பட்ட மின்னணு அகராதிகள் பாரம்பரிய அச்ச வடிவத்தை விட பல நன்மைகளை வழங்குகின்றன, இருப்பினும் அனைத்து மின்-அகராதிகளும் மின்னணு ஊடகம் வழங்கும் அம்சங்களை உண்மையில் இணைக்காது. மரபு ஒலிப்பு எழுத்தாக்கத்தைத் (phonetic transcription) தவிர, அகராதி பயனர்கள் (குறிப்பாக மொழியைக் கற்றுக்கொள்பவர்கள்) ஒரு சொந்த மொழிபேசுபவரால் (native speaker) உச்சரிக்கப்படும் தலைபுச்சொல்லை (ஒருவேளை எடுத்துக்காட்டு வாக்கியங்கள்) கேட்க முடியும். தேடப்பட்ட சொல்லைத் தவறாக எழுத்துக்கூட்டுவது தோல்வியை விளைவிப்பதில்லை, ஏனெனில் மின்னணு அகராதிகள் உண்மையில் அச்சொல் என்ன அர்த்தத்தைத் தருகிறது என்பதை யூகிக்க முயற்சிக்கும் (Lew & Mitton, 2011, 2013).

காகித அகராதிகளின் அறியப்பட்ட சிக்கல் (அகரவரிசை எழுத்து முறைகளைக் கொண்ட மொழிகளுக்கு) அவை மரபாக ஒற்றை எழுத்துச் சொற்களைச் சுற்றி ஒழுங்கமைக்கப்படுகின்றன. பல்சொல் அலகுகளை (multi-word lexical units) (மரபுச்சொற்கள் போன்றவை) கண்டுபிடிக்க முயற்சிக்கும் அகராதிப் பயனர்களுக்கு இது குறிப்பிட்ட சிரமங்களை உருவாக்குகிறது. மின் அகராதியின் (electronic dictionary) நன்கு வடிவமைக்கப்பட்ட இடைமுகம், பல்சொல் அலகு என்பது அகராதி

பயனருக்குத் தெரியாத சந்தர்ப்பங்களில் கூட, எந்தத் தலைச்சொல்லின் கீழ் வெளிப்பாடு உள்ளமைக்கப்பட்டுள்ளது என்பதை அறியாமலேயே பல்சொல் அலகைக் கண்டறிவதை சாத்தியமாக்கும்; இது மொழி கற்பவர்கள் வெளிநாட்டு மொழியில் உள்ள உரையைப் புரிந்துகொள்வதில் சிரமப்படுவதைப் போன்றது.

மின்னணு வடிவத்தில் நூல்களைப் படிக்கும்போது, வாசிப்புக்கான மென்பொருளும் அகராதியும் ஒன்றுக்கொன்று "பேச" முடிந்தவுடன் அகராதி ஆலோசனையை எளிதாக்கலாம். எடுத்துக்காட்டாக, மின் புத்தகம் படிக்கும் சாதனத்தில், பயனர் செய்ய வேண்டியதெல்லாம், பிரச்சனைக்குரியது என்று அவர்கள் நம்பும் சொல்லைத் தட்டினால் மட்டுமே, இந்தச் சொல் தானாகவே பார்க்கப்பட வேண்டும். சிறந்த வகையில், ஸ்மார்ட் அகராதியானது (smart dictionary) பல்சொல் வெளிப்பாடுகளுக்கான (multi-word expressions) சான்றுகளுக்காகவும் மற்றும் தேடல் வார்த்தை குறிப்பிட்ட அர்த்தத்திற்காகப் பயன்படுத்தப்பட வாய்ப்புள்ளது என்ற தடையங்களுக்காகவும் உரைச் சூழலை ஆய்வு செய்ய வேண்டும். பின்னர், வழங்கப்பட்ட பதிவு பொருத்தமற்றதாக இருக்கக்கூடிய தகவலை அடக்குவதன் மூலமும் மற்றும் இந்த குறிப்பிட்ட புரிதல் சிக்கலில் மதிப்புமிக்க தரவை தேர்ந்தெடுத்து வழங்குதல் மூலமும் இந்த கண்டுபிடிப்புகளின் விளைவுகளை பிரதிபலிக்க வேண்டும்.

அச்சிடப்பட்ட அகராதிகளைப் போலவே மின்னணு அகராதிகள் ஒற்றை நிலையான பார்வைக்குக் கட்டுப்படுத்தப்பட வேண்டியதில்லை. அகராதிசார் தரவுகளின் வழங்கல், பயன்பாடுகள் மற்றும் பயனர்களைப் பொறுத்துச் சரிசெய்யப்படலாம். எளிமையான நேர்வில், ஒரு வெளிநாட்டு மொழியில் ஒரு உரையைப் படிக்கும் போது ஒரு அகராதியை ஆலோசிப்பவருக்கு இலக்கண நிரப்புதல் (grammatical complementation), சொல்லடிவகைப்பாட்டு நடத்தை (collocational

behavior) அல்லது ஒருபொருள்பன்மொழிகள் பற்றிய விரிவான தகவல்கள் மிதமிஞ்சியவை. ஆனால் அதே விரிவான தரவு ஒரு கட்டுரை எழுதும் ஒருவருக்கு மிகவும் பயனுள்ளதாக இருக்கும்.

மின்னணு ஊடகத்தின் (electronic media) சேமிப்பு இடம் அச்சிடப்பட்ட தொகுதிகளில் முடிந்ததை விட அதிகமான அகராதித் தரவுகளை உள்ளடக்குவதைச் சாத்தியமாக்குகிறது. கிடைக்கக்கூடிய எல்லாத் தரவும் எல்லா நேரங்களிலும் பயனருக்கு வழங்கப்பட வேண்டும் என்று இது அர்த்தப்படுத்துவதில்லை, ஏனெனில் அவ்வாறு செய்வது பெரும்பாலும் தகவல் சுமைக்கு வழிவகுக்கும், மேலும் சொரற்பதிவு (entry) பயனுள்ளதாக இருக்காது. எவ்வாறாயினும், அந்த நேரத்தில் என்ன தரவு வழங்கப்பட வேண்டும் என்பதில் சரியான வகையான கட்டுப்பாட்டுடன், செழிப்பான தரவு சேமிப்பகங்களை அணுகுவது அச்சிடப்பட்ட அகராதிகளை விட குறிப்பிடத்தக்க மேம்பாடுகளை வழங்கக்கூடும். எடுத்துக்காட்டாக, கூடுதல் எடுத்துக்காட்டு வாக்கியங்கள் சேர்க்கப்படலாம். மிகவும் மேம்பட்ட அகராதி பயனர்களுக்கு அத்தகைய எடுத்துக்காட்டு வங்கி, முழு அம்சம் கொண்ட உரைத் தரவுத்தொகுதி வடிவத்தை எடுக்கலாம்.

மின்னணு காட்சிகள் அச்சிடப்பட்ட அகராதிகளிலிருந்து பெறப்பட்ட வழக்கமான உரைக்கும் (மற்றும் சில நேரங்களில் அவ்வப்போது வரும் படம்) அப்பால் ஊடக உள்ளடக்கத்திற்கு இடமளிக்கும். முழு வண்ணப் புகைப்படங்கள் போன்ற வரைகலைகளை/கிராபிக்ஸ்களை வழக்கமாகச் சேர்ப்பது அதிக வெளியீட்டுச் செலவுகளாக மாறாது, மேலும் இவை விருப்பமான கூறுகளாக வழங்கப்படும் வரை, அவை மதிப்புமிக்க இடத்தைப் பயன்படுத்தாது. வீடியோ/காணொலி மற்றும் அனிமேஷன்/அசைவூட்டல் மற்றொரு சாத்தியம், இருப்பினும் கிடைக்கக்கூடிய சான்றுகள் அவை பயனருக்குப் பயனளிக்காது என்று கூறுகின்றன (சுன் & பிளாஸ்,

1996; லூ & டோரோஸ்ஸெவ்ஸ்கா/Chun & Plass, 1996; Lew & Doroszewska, 2009). மின்னணு/எலக்ட்ரானிக் சாதனத்தில் ஒலித் திறன் இருந்தால், தலைச்சொற்களின் முன்னரே பதிவுசெய்யப்பட்ட அல்லது தொகுக்கப்பட்ட ஒலிப் பொருள் (audio material) மற்றும் சாத்தியமாக எடுத்துக்காட்டு வாக்கியங்கள் மற்றும் வரையறை விளக்கங்களை அகராதி பயன்படுத்தலாம். சில சொற்கள், குறிப்பாக இசைக்கருவிகளைக் குறிக்கும், சிறப்பியல்பு ஒலிகளுடன் தொடர்புடையவை, மேலும் இவற்றின் பதிவுகள் பயனர்களின் புரிதலை மேம்படுத்தும், மேலும் நீண்ட கால விளைவாக, இந்தச் சொற்றொகைப் பொருட்களைத் (vocabulary items) தக்கவைத்துக்கொள்ளும்.

நவீன இணைய தொழில்நுட்பம் (internet technology) குறிப்பாக உள்ளடக்கத்தை இணைப்பதையும் உட்பொதிப்பதையும் எளிதாக்குகிறது. இது பன்மொழிகளுக்கான (எ.கா. <http://dict.cc/>) புதுமையான சொல்சார் ஆதாரங்களை உருவாக்குவதற்கான சாத்தியக்கூறுகளைக் கொண்டிருந்தாலும், சில சமயங்களில் திரட்டி தளங்கள் (aggregator sites) (Lew 2011) என்று அழைக்கப்படுவதைத் தயாரிப்பதற்கு இந்த விருப்பம் தவறாகப் பயன்படுத்தப்படுகிறது, இது தரத்தில் உண்மையான அக்கறை இல்லாமல் ஆதாரங்கள் பலவற்றிலிருந்து உள்ளடக்கத்தை இழுக்கிறது. நிபுணர்கள் அல்லாத பயனர்கள், சரியான வழிகாட்டுதல் இல்லாமல் தங்கள் உண்மையான மதிப்பை மதிப்பிட முடியாமல், பொதுமைப்படுத்தப்பட்ட பொருண்மைக்களப் பெயர்கள் அல்லது உள்ளடக்கிய சொற்கள் அல்லது மொழிகளின் எண்ணிக்கையின் உயர்த்தப்பட்ட கூற்றுகளால் ஈர்க்கப்படலாம். இதற்கு நேர்மாறாக, மிகவும் செயலில் உள்ள பயனர்கள், Web 2.0 தொழில்நுட்பத்தைப் பயன்படுத்தி, அகராதியை தங்கள் கைகளில் எடுத்துக்கொள்கிறார்கள். இத்தகைய பயனர்

உருவாக்கிய உள்ளடக்கம், சிறப்பு அகராதியியலின் (Lew 2013) பகுதியில் மிகவும் பயனுள்ளதாக இருக்கும்.

நவீன தொழில்நுட்பமானது பல்வேறு வகையான அகராதிகளுக்கும் அகராதிகளுக்கும் சொல் அடிப்படையிலான கருவிகளுக்கும் இடையே உள்ள பாரம்பரிய/மரபு வேறுபாடுகளை மங்கலாக்குகிறது. ஏற்கனவே மேலே கொடுக்கப்பட்டுள்ள இரண்டு எடுத்துக்காட்டுகளைக் குறிப்பிடலாம்: ForbetterEnglish.com என்பது தானாக உருவாக்கப்பட்ட சொல்லடிவகைப்பாடுகளின் (collocations) அகராதியாகும், தரவுத்தொகுதியிலிருந்து தானாகவே தேர்ந்தெடுக்கப்பட்ட எடுத்துக்காட்டுகளுடன் விளக்கப்பட்டுள்ளது; Just the Word/ஜஸ்ட் த வேர்ட் (<http://www.just-the-word.com/>) என்பது ஒரு சுவாரஸ்யமான சொல்சார் விவரக்குறிப்புக் கருவியாகும், இது தரவுத்தொகுதிச் சான்றுகளின் அடிப்படையில் சொந்தமொழி அல்லாத சொல்லடிவகைப்பாட்டுத் தேர்வுகளைச் சரிசெய்யும் திறன் கொண்டது. இன்று, மொழி கற்பவர்கள் அல்லது மொழிபெயர்ப்பாளர்களுக்கான பெரிய மென்பொருள் தொகுப்புகளின் ஒரு பகுதியாக அகராதிகள் ஒருங்கிணைக்கப்படலாம், அதாவது எழுத்து உதவியாளர்கள் (writing assistants) (எழுத்தும் பணிகளில் ஆதரவை வழங்க வடிவமைக்கப்பட்ட மென்பொருள்) போன்று. இந்தப் போக்கு தொடரும், எதிர்காலத்தின் அகராதிகள் நமக்குத் தெரிந்தவற்றிலிருந்து வேறுபட்டதாக இருக்கும்.

முடிவுரை

கணிப்பொறி வழி அகராதி தயாரிக்கும் படிநிலைகளில் கணிப்பொறி படுத்த இயலும் நிலைகளையும், முறைகளையும் சுருக்கமாகப் பார்த்தோம். இம்முறைகளில் அகராதியில் மொழி ஒரு தடையில்லை. தமிழ் அகராதியோ, ஆங்கில அகராதியோ, பிறமொழி அகராதிகளையோ மேற்கண்ட முறைகளைப் பின்பற்றி உருவாக்கலாம்.

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

அந்தந்த மொழிகளில் அகர வரிசைப்படுத்தப் படும் மென்பொருள்களை வடிவமைக்க வேண்டும்.

மொழிபெயர்ப்பு ஒரு கலை. அது ஒரு படைப்புச் செயல்பாடும் கூட. ஆனால், இன்றைக்கு அது தேவையாகிவிட்டது. இணையத்தில் கொட்டிக் கிடக்கும் தகவல்களை நாம் எளிதில் படித்துத் தெரிந்துகொள்ள மொழிபெயர்ப்பை நாம் தெரிந்திருக்க வேண்டியதிருக்கிறது. அதனால் ஒரு அகராதியும் (Dictionary) நமக்கு அவசியமாகிறது. அதுபோல் படிக்கும் மாணவர்களுக்கும் தங்கள் பாடம் தொடர்பாகச் சில விஷயங்களை விளங்கிக்கொள்ள அகராதி அவசியம். இதற்காக முன்புபோல் கனத்த ஆங்கில-தமிழ் அகராதியைத் விரித்துப் பார்த்து பென்சிலில் குறித்துக்கொள்ள வேண்டியதில்லை. இப்போது இணையத்தில் பல அகராதிகள் கிடைக்கின்றன. வேண்டிய சொல்லை இடுகையில் இட்டால் போதும், அது தொடர்பான பொருட்கள் பல வந்து திரையில் விழும். இணையத்தில் ஆங்கிலம்-ஆங்கிலத்துக்கான அகராதிகள் பல கிடைக்கின்றன. ஆனால், அதை ஒப்பிடும்போது தமிழ்-ஆங்கில அகராதிகள் குறைவு. கூகுளின் இலவசச் சேவையான கூகுள் மொழிபெயர்ப்பாளர்/டிரான்ஸ்லேட்டர் அதிகமாகப் பயன்படுத்தப்படும் அகராதியாக இருக்கிறது. இதில் தமிழ் மட்டுமல்லாது நூற்றுக்கும் மேற்பட்ட மொழிகளுக்கான அகராதியாக அது இருக்கிறது. ஆனால் கூகுள் மொழிபெயர்ப்பாளர்/டிரான்ஸ்லேட்டர் அவ்வளவு சரியான பொருளை அளிப்பதில்லை. ஓரளவுக்குத்தான் அது பொருள் தரும். இந்நிலையில் நம்பகமான சில ஆங்கிலம்- தமிழ் அகராதிகளைப் பார்க்கலாம்.

3. சொற்களஞ்சியமும் சொற்களஞ்சிய உருவாக்கமும்

ஒரு சொற்களஞ்சியம் (thesaurus - plural thesauri or thesauruses), சில சமயங்களில் ஒருபொருள்பன்மொழி அகராதி (synonym dictionary) என்று அழைக்கப்படுகிறது, இது ஒரு நோக்கீட்டுப் படைப்பாகும். இது பொதுவாக ஒத்த பொருளுடைய சொற்கள் மற்றும் தொடர்களின் பட்டியல்கள் அடங்கிய ஏடு என வரையறை விளக்கம் செய்யப்படுகின்றது; இது சொற்களை அவற்றின் அர்த்தங்களின்படி வரிசைப்படுத்துகிறது, சில நேரங்களில் பரந்த மற்றும் குறுகிய சொற்களின் படிநிலையாக, சில நேரங்களில் வெறுமனே பட்டியல்களாக. ஒருபொருள்பன்மொழிகள் மற்றும் எதிர்ச்சொற்கள். ஒரு கருத்தை வெளிப்படுத்த சிறந்த வார்த்தையைக் கண்டறிய உதவுவதற்கு அவை பெரும்பாலும் எழுத்தாளர்களால் பயன்படுத்தப்படுகின்றன:

ஒருபொருள்பன்மொழி அகராதிகளுக்கு நீண்ட வரலாறு உண்டு. 1852-ஆம் ஆண்டில் பீட்டர் மார்க் ரோஜெட் (Peter Mark Roget) தனது ரோஜெட்ஸ் தெசரஸுக்கு (Roget's Thesaurus) 'thesaurus' என்ற வார்த்தையைப் பயன்படுத்தினார். "thesauri" என்று அழைக்கப்படும் சில படைப்புகள், ரோஜெட்டின் தெசரஸ் போன்றவை, கருத்துகளின் படிநிலை ஹைப்பர்னிமிக் வகைபிரிப்பில் உள்ள குழு வார்த்தைகள், மற்றவை அகரவரிசைப்படி அல்லது வேறு வழிகளில் ஒழுங்கமைக்கப்பட்டுள்ளன. பெரும்பாலான சொற்களஞ்சியம் வரையறைகளை உள்ளடக்குவதில்லை, ஆனால் பல அகராதிகளில் ஒருபொருள்பன்மொழிகளின் பட்டியல்கள் உள்ளன. சில சொற்களஞ்சியம் மற்றும் அகராதி ஒத்த குறிப்புகள் ஒரே மாதிரியான சொற்களுக்கு இடையே உள்ள வேறுபாடுகளை வகைப்படுத்துகின்றன, அவற்றின் "குறிப்புகள் மற்றும் அர்த்தத்தின் மாறுபட்ட நிழல்கள்" பற்றிய குறிப்புகள். சில ஒருபொருள்பன்மொழி அகராதிகள் முதன்மையாக பொருள் மற்றும் பயன்பாடு மூலம்

ஒருபொருள்பன்மொழிகளை வேறுபடுத்துவதில் அக்கறை கொண்டுள்ளன. ஃபோலரின் நவீன ஆங்கில பயன்பாட்டின் அகராதி அல்லது கார்னரின் நவீன ஆங்கில பயன்பாடு போன்ற பயன்பாட்டுக் கையேடுகள் பெரும்பாலும் ஒருபொருள்பன்மொழிகளின் பொருத்தமான பயன்பாட்டைப் பரிந்துரைக்கின்றன

எழுத்தாளர்கள் சில சமயங்களில் சொற்களைத் திரும்பத் திரும்பச் சொல்வதைத் தவிர்க்க ஒருபொருள்பன்மொழிகளைப் பயன்படுத்துகின்றனர்; இது நேர்த்தியான மாறுபாடு; இது பெரும்பாலும் பயன்பாட்டுக் கையேடுகளால் விமர்சிக்கப்படுகிறது: "எழுத்தாளர்கள் சில சமயங்களில் தங்கள் சொற்றொகையை மாற்றுவதற்கு மட்டுமல்லாமல், அவற்றை அதிகமாக அலங்கரிக்கவும் பயன்படுத்துகிறார்கள்".

விரிவான பொருளில் சொற்களின் தொகுப்பைச் சொற்களஞ்சியம் (thesaurus) எனலாம். சொற்களஞ்சியம் என்பதை பொருட்புல அகராதி என்றும் குறிப்பிடுவர். சொற்களஞ்சியம் கருத்துருக்களிலிருந்து அல்லது பொருண்மைகளிலிருந்து (அர்த்தங்களிலிருந்து) சொற்களை அடைய உதவும். சொற்களஞ்சியம் என்பதன் விளக்கம் பின்வரும் ஜோன்ஸின் (Jones, 1986:201) மேற்கோள் மூலம் தெளிவாகும்: "ஒரு சொற்களஞ்சியத்தில் ஒருபொருள் பன்மொழிகள் இருக்க வேண்டிய கட்டாயம் இல்லை. மற்றும் ஒருபொருள் பன்மொழி அகராதி சொற்களஞ்சியமாக அமையத் தேவையில்லை. விரிவான பொருளில் சொற்களஞ்சியம் என்பது சொற்களைக் கருத்துருக்கள் (Concepts), தலைப்புகள் (topics) அல்லது பாடப்பொருள்கள் (Subjects) ஆகியவற்றின் அடிப்படையில் பாகுபாடு செய்வதாகும்; ஒரு வகுப்பில் தரப்பட்டுள்ள சில சொற்கள் ஒருபொருள் பன்மொழிகளாகும் என்ற செய்தி நிச்சயமாகத் தேவையற்ற ஒன்றாகும். மாறாக ஒருபொருள் பன்மொழி அகராதி ரொஜெஸ்டின் சொற்களஞ்சியத்தில் (Rogest's Thesaurus) தரப்பட்டுள்ள கருத்துருப் பாகுபாட்டைத் தராமல் குழுமங்களாக வரும் சொற்களைத்தான் பகுத்தளிக்கும். இது ஒரு பக்குவமற்ற வேற்றுமைப்படுத்தலாகும். சொற்களஞ்சியத்தில் ஒருபொருள் பன்மொழிகள் அடிக்கடி காணப்படும். ஒருபொருள் பன்மொழி அகராதிகளில் உள்ள குழுமங்கள் பெரிதாகவும் ஒருபொருள் பன்மொழியத்தின் பரந்த விளக்கத்திற்கு மாதிரியாகவும் அமைந்துள்ள

குழுமங்களை இணைக்கும் குறுக்கு நோக்கீட்டுகளையும் கொண்டிருக்குமானால் அவை சொற்களஞ்சியத் தன்மை பெறலாம்.”

சொற்களஞ்சியம் பற்றிய விரிவான தகவல்கள் இராசேந்திரன் அவர்களால் (இராசேந்திரன், 2001) தரப்பட்டுள்ளன. சொற்களஞ்சியம் என்பது சொற்களை வரிசைப்படுத்தித் தருவது என்ற குறுகிய நோக்கில் அமையாமல் சொற்களின் வலைபோல் பின்னிக்கிடக்கும் பொருண்மை உறவுகளையும் சொல் உறவுகளையும் வெளிப்படுத்த வேண்டும் என்ற நோக்கில் அமைப்பாக்கம் செய்யப்படவேண்டும்.

சொற்களஞ்சியத்தில் சொற்றொகையைச் சார்ந்த பெயர், வினை, பெயரடை, வினையடை, பின்னருபுகள், இணைப்பான்கள் மற்றும் பிற செயல்பாட்டுச் சொற்கள் (functional words) தரவுகளாச் சேகரிக்கப்பட்டு ஒரு வகைப்பாட்டியல் (taxonomy) வடிவத்தில் தர முயற்சிக்கப்பட்டுள்ளது.

2.1. சொற்களஞ்சியத்தின் முன்னோடிகள்

சொற்களஞ்சியம் தோன்றுவதற்கு முன்னர் சொற்களுக்கான பொருளை அறிந்துகொள்ள நிகண்டு, அகராதி போன்ற நூல்கள் உருவாக்கப்பட்டன. அவை சொற்களஞ்சியத்திற்கு முன்னோடிகளாக அமைகின்றன.

ஒரு மொழியின் வளர்ச்சியையும் அதன் தனித்தன்மையையும் தெரிவிக்கும் துல்லிய அளவுகோலாகச் சொற்றொகை அமைகிறது. சொற்களைப் பற்றிய ஆய்வும் சொற்களஞ்சிய உருவாக்கமும் இலக்கண உருவாக்கக் காலத்திலேயே தொடங்கப்பட்டன. தொல்காப்பியத்தில் சொற்களஞ்சியம் இலக்கணத்தின் பகுதியாகப் பல இடங்களில் சுட்டப்பெறுகிறது. இதற்குப் பிறகு சொற்றொகுதி தேவையை உணர்ந்து நிகண்டுகள் உருவாக்கப்பட்டன.

நிகண்டுகள்

முதன்முதலில் நிகண்டு என்பது வடமொழியில் வேதத்திற்கு அங்கமாய், வைதிகச் சொற்களின் பொருள்களை உணர்த்தும் கருவி நூலுக்கு மட்டும் சிறப்பாக வழங்கப்பட்டது. பின்னர் நாளடைவில் சொற்பொருள் உணர்த்தும் நூல்களுக்கு எல்லாம் இப்பெயர் வழங்கத் தலைப்படலாயிற்று. நிகண்டு என்பது தமிழில் உரிச்சொல், உரிச்சொற்பொருள் என்று வழங்கப்பட்டது. மிகப் பழமையான நிகண்டு

நூல் திவாகர நிகண்டு ஆகும். பின்னர் வந்த நிகண்டுகள் அனைத்தும் திவாகரத்தை அடிப்படையாக அமைந்தது.

நிகண்டுகள் நூற்பா வடிவில் அமைந்திருந்தன. மனப்பாடம் செய்வது கடினமாக இருந்தது. எனவே அதனை எளிமைப்படுத்தும் நோக்கில் வெண்பா, ஆசிரியப்பா, கலிப்பா, விருத்தப்பா போன்ற பாவகைகளைக் கொண்டு நிகண்டுகள் பின்னர் உருவாக்கப்பட்டன. இதனால் சொற்களுக்குப் பொருளறிவது எளிமையாக இருக்கிறது.

திவாகர நிகண்டு

திவார நிகண்டின் சொற்றொகைப் பாகுபாடு

திவாகர நிகண்டில் சொற்றொகைப் பன்னிரண்டு தொகுதிகளாகப் பிரிக்கப்பட்டுள்ளன.

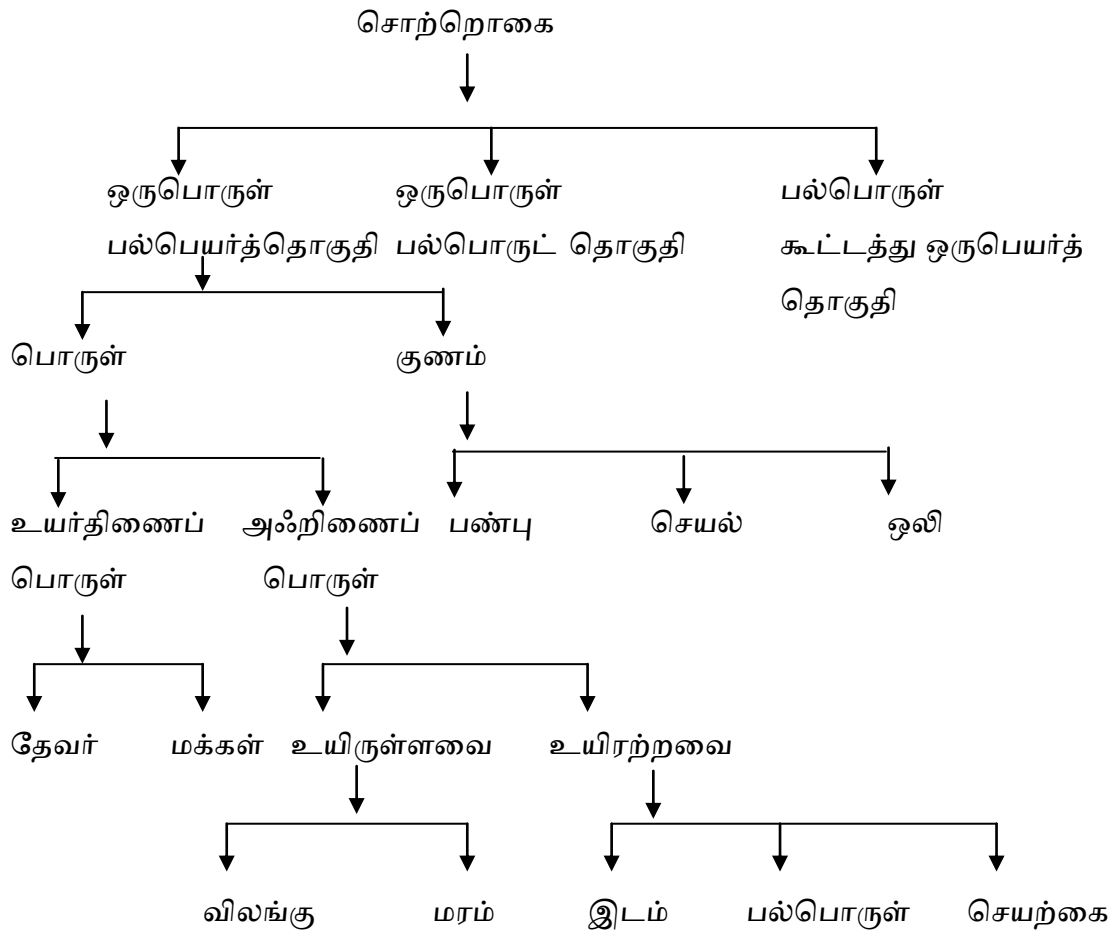
1. தெய்வப்பெயர்த் தொகுதி: சிவன், திருமால் முதலிய தெய்வப்பெயர்கள், நாள் கோள்களின் பெயர், காலப்பெயர், பருவப்பெயர், மழை, மேகம் முதலியவற்றின் பெயர்கள்.
2. மக்கட் பெயர்த் தொகுதி: அருந்தவர், அறிஞர், பரிவாரங்கள், தொழிலாளர்கள், கீழ்மக்கள், பெண்கள், நானிலமக்கள், உறவினர், படைகள், உடலுறுப்புகள் முதலியவற்றின் பெயர்கள்
3. விலங்கினப் பெயர்த் தொகுதி: விலங்கு, பறவை, ஊர்வன, நீர்வாழ்வன முதலியவற்றின் பெயர்கள், அவற்றின் உறுப்புப்பெயர்கள்
4. மரப்பெயர்த் தொகுதி: மரம், செடி, கொடி, பூ, காய், கனி, தானியவகை முதலியவற்றின் பெயர்கள்.
5. இடப்பெயர்த் தொகுதி: உலகம், நாடு, ஊர், மனை, மலை, ஆறு, நீர்நிலை, முதலியவற்றின் பெயர்கள்.
6. பல்பொருட் பெயர்த் தொகுதி: பொன், மணி, விரைப்பொருள், உணவுவகை முதலியவற்றின் பெயர்கள்.
7. செயற்கை வடிவப் பெயர்த் தொகுதி: போக்கருவிகள், அணிகலன் வகை, உடைவகை, ஊர்தி வகை, விளையாட்டுக் கருவிகள் முதலியவற்றின் பெயர்கள்.
8. பண்புபற்றிய பெயர்த்தொகுதி: காட்சிப்பொருள், உணவிற்குப் புலப்படும் உண்மைப் பொருள் இவற்றின் குணத்தொடர்பு பற்றிய பெயர்கள்
9. செயல் பற்றிய பெயர்த்தொகுதி: வினைப்பெயர்கள்

10. ஒலிபற்றிய பெயர்த்தொகுதி: சொல்வகை, பாவகை, யாழ்வகை, இசை வகை முதலியன குறித்த பெயர்கள்.

11. ஒருசொல் பல்பொருட் பெயர்த்தொகுதி: ஒருசொல்லுக்குரிய பல்வேறு பொருள்கள்

12. பல்பொருட்கூட்டத்து ஒருபெயர்த்த தொகுதி: இருசுடர், எண்குணம் என்றாற் போன்ற தொகைப் பொருட்கள்

இச்சொற்றொகைப் பாகுபாட்டை பின்வருமாறு கிளைப்படமாகத் தரலாம்.



முப்பொருட் பகுதிகள்

திவாகரத்தின் பன்னிரண்டு தொகுதிகளை பின்வரும் மூன்று பெரும் பிரிவுகளுக்குள் அடக்கலாம்:

- ஒருபொருள் பல்பெயர்த்த தொகுதி
- ஒருபொருள் பல்பொருட் தொகுதி

- பல்பொருள் கூட்டத்து ஒருபெயர்த் தொகுதி

நாமதீப நிகண்டு

நிகண்டு சொற்றொகை வகைப்பாடுகளில் நாமதீப நிகண்டில் காணப்படும் சொற்றொகை வகைப்பாடு குறிப்பிடத்தக்க விதத்தில் அமைந்துள்ளது. நாமதீப நிகண்டின் சொற்றொகைப் பகுப்பு பின்வருமாறு அமைந்துள்ளது.

1. உயர்திணைப்படலம்

- 1.1. தெய்வ வர்க்கம்
- 1.2. மானிட வர்க்கம்

2. அஃறிணையுயிர்ப்படலம்

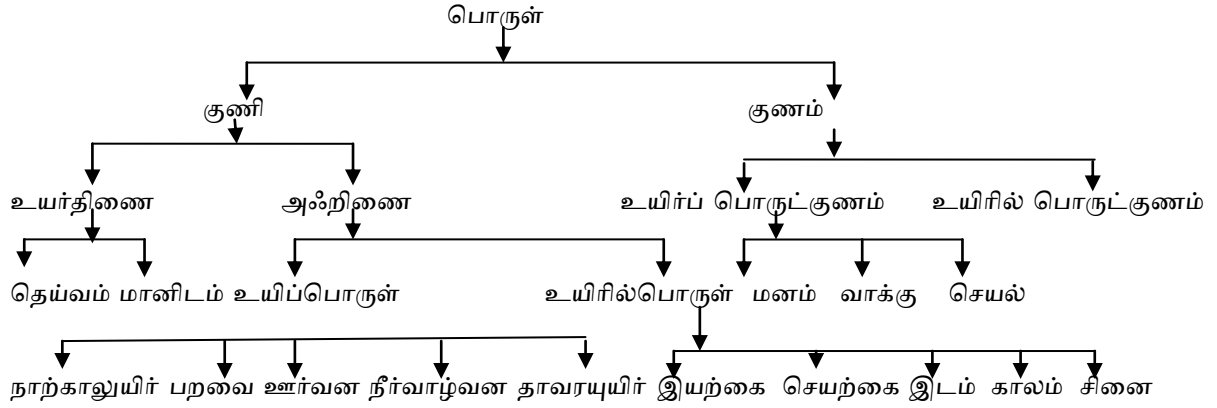
- 2.1. நாலுகாலுயிர்ப் பொருள் வர்க்கம்
- 2.2. பறவையுயிர்ப் பொருள் வர்க்கம்
- 2.3. ஊருயிர்ப் பொருள் வர்க்கம்
- 2.4. நீர்வாழ்வுயிர்ப் பொருள் வர்க்கம்
- 2.5. தாவரவுயிர்ப் பொருள் வர்க்கம்

3. அஃறிணையுயிரில் பொருட்படலம்

- 3.1. இயற்கைப் பொருள் வர்க்கம்
- 3.2. செயற்கைப் பொருள் வர்க்கம்
- 3.3. இடவர்க்கம்
- 3.4. காலவர்க்கம்
- 3.5. சினைவர்க்கம்

4. குணநாமப்படலம்

- 4.1. உயிர்ப் பொருள் மனக்குணவர்க்கம்
- 4.2. உயிர்ப்பொருள் வாக்குக் குணவர்க்கம்
- 4.3. உயிர்ப்பொருள் செயற்குணவர்க்கம்
- 4.4. உயிரில்பொருள் குணவர்க்கம்



ராஜெஸ்ட் தெசாரஸ் அடிப்படையிலான சொற்களஞ்சியம் (ROGETS THESAURUS)

ராஜெஸ்ட் என்பவர் தனது தெசாரஸ்சை (சொற்களஞ்சியத்தை) 1852-இல் வெளியிட்டார். இன்றைய ஆங்கில மொழிக்கான தெசாரஸில் பெரும்பாலானவை ராஜெஸ்ட் தெசாரஸ்சை அடிப்படையாக்கொண்டு உருவாக்கப்பட்டவை. பின்வரும் வகைப்பாட்டுத்திட்டம் அவருடைய வகைப்பாட்டின் அடிப்படையில் அமைந்துள்ளது. வகைப்பாட்டுத் திட்டம் (Plan of Classification)

Class	Section	Sub section
ABSTRACT RELATIONS	I. Existence	
	II. Relation	
	III. Quantity	
	IV. Order	
	V. Number	
	VI. Time	
	VII. Change	
	VIII. Causation	
SPACE	I. In General	
	II. Dimension	
	III. Form	
	IV. Motion	
MATTER	I. In General	
	II. Inorganic	1. Solids

		2. Fluids
	III. Organic	1. Vitality
		2. Sensation
INTELLECT	I. Formation of Ideas	
	II. Communication of ideas	
VOLITION	I. Individual	
	II. Intersocial	1. In General
		2. Possessive relations
AFFECTIONS	I. In General	
	II. Personal	
	III. Sympathetic	
	IV. Moral	
	V. Religious	

அகராதிகள்

அகராதிகள் ஒரு மொழியின் சொற்களை அகரவரிசைப்படுத்தி பொருள் கூறுகின்றன. சொற்களை அகரவரிசையில் அமைப்பதன் இன்றியமையாமை நிகண்டுகளின் காலத்திலேயே உணரப்பெற்றது. குறிப்புகள் நோக்கீட்டிற்கு உதவும் வண்ணம் நிகண்டுகளை மாற்றி அமைக்கும் முயற்சிகளை நிகண்டுகளின் வளர்ச்சி வரலாற்றில் நின்றும் அறியலாம். முதலில் சொற்கள் எதுகை மோனை அமைய வரிசைப்படுத்தப்படாததும், பின்னர் ககர எதுகை, சகர எதுகை அமையுமாறு வரிசைப்படுத்தப்பட்டதையும் அகராதியை நோக்கிய தொடக்க முயற்சிகள் எனலாம். பத்தொன்பதாம் நூற்றாண்டும் இருபதாம் நூற்றாண்டும் தமிழ் அகராதி உருவாக்கத்தின் வேகமான முன்னேற்றத்தைக் காட்டுகின்றன. சென்னைப் பல்கலைக் கழகத்தால் வெளியிடப்பட்ட Tamil Lexicon என்ற தமிழ்-ஆங்கில-தமிழ் அகராதி இன்றைய அகராதியியலாருக்குக் கருவூலமாக அமைகின்றது. 1992-இல் முதற் பதிப்பாக வெளியிடப்பட்டு 2008-இல் இரண்டாம் பதிப்பாக வெளிவந்துள்ள தமிழ்-தமிழ்-ஆங்கில அகராதியான 'க்ரியாவின் தற்காலத் தமிழ் அகராதி' தமிழ்மொழிசார் அகராதியியலின் உச்சகட்ட வளர்ச்சியைப் பறைசாற்றி நிற்கின்றது. க்ரியா அகராதியின்

=====

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

விரிவாக்கப்பட்ட பதிப்புக்கான சொல்வங்கி முதல் பதிப்பின் 18 லட்சம் சொற்கள் என்ற அளவிலிருந்து 75 லட்சம் என்ற அளவை எட்டியுள்ளது (க்ரியாவின் தற்காலத் தமிழகராதி, 2008:xxvii).

2.2. சொற்களஞ்சியத்தின் தேவை

தமிழ்ச் சொற்களஞ்சியம் உருவாக்க நிகண்டுகள் முன்னோடிகளாக அமைந்தன. ஆனால் நிகண்டுகளின் சொற்பாடுபாடுகள் மேல்நிலையில் நிறைவு தருவனவாக இருந்தாலும் பெருந்தொகுதிகளில் அடங்கும் சிறு பாடுபாடுகள் சரியான முறையில் வரையறுக்கப்படவில்லை. அதிகமாக ஒரு தலைச்சொல்லின் கீழ் ஒருபொருள் பன்மொழிகளைத் தருவதே வழக்கமாயுள்ளது. பெருந்தொகுதிகள் மேலும் சிறுதொகுதிகளாகப் படிநிலை அமைப்பில் பகுக்கப்படவில்லை. நிகண்டுகளில் எதுகை, மோனைக்கு தந்த இன்றியமையாமை பொருண்மைக்குத் தரப்படவில்லை.

தமிழில் வழக்கிலுள்ள எல்லாச் சொற்களும் நிகண்டுகளில் தரப்படவில்லை. மேலும் தற்காலத் தமிழுக்குத் தேவையான பல சொற்கள் நிகண்டுகளில் இல்லை. நிகண்டுகளில் தந்துள்ள பல சொற்கள் தற்காலத் தமிழில் வழக்கிறந்தவைகளாக உள்ளன. இந்த குறைபாடுகளைப் போக்கும் வகையில் தற்காலத் தமிழுக்குச் சொற்களஞ்சியம் உருவாக்கப்பட்டுள்ளது (இராசேந்திரன் 2001).

நிகண்டுகள் தமிழ்ச் சொற்றொகையை ஒருபொருள் பன்மொழிப் பெயர்களாகவும் ஒருசொல் பல்பொருள் பெயர்களாகவும் தொகைப்பெயர்களாகவும் பகுத்துச் சூத்திர வடிவத்தில் தந்து அதன் மூலம் சொற்களை மனைஞ் செய்து சொல் வளத்தைப் பெருக்க உதவும் மனைக் கருவிகளாகச் செயல்படுகின்றன. அவை முற்காலத் தமிழைத் தழுவி உருவாக்கப்பட்டவை. எனவே தற்காலத் தமிழுக்கு நிகண்டு போல் மனைக் கருவியாக அமையாமல் நோக்கீட்டுக் கருவியாக அமையும் சொற்களஞ்சியம் ஒன்று உருவாக்கும் தேவை ஏற்பட்டுள்ளது.

5. சொற்களஞ்சியத்தின் பின்னணியில் அமையும் பொருண்மைக் கோட்பாடுகள்

இராசேந்திரனின் தற்காலத் தமிழ்ச் சொற்களஞ்சியம் (இராசேந்திரன் 2001) நைடா (Nida, E.A. 1975a, 1975b, 1975c) லையான்ஸ் (Lions 1963, 1977), குரூஸ் (Cruse

1986, 200) ஆகியோர்களின் பொருண்மைக் கோட்பாடு மற்றும் கொள்கைகள் அடிப்படையில் உருவாக்கப்பட்டுள்ளது.

சொற்களின் பொருண்மை உறவுகள்

சொற்களின் பொருண்மையை இருவகை உறவுகள் அடிப்படையில் அணுகலாம்:

1. உறுப்பமைவு உறவு அடிப்படையில் (by syntagmatic relation)
2. அடுக்கு உறவு அடிப்படையில் (by paradigmatic relation)

உறுப்பமைவு உறவு

ஒரு வாக்கியத்தில் ஒன்றையொன்று தொடர்ந்து வரும் சொற்களுக்கு இடையே உள்ள உறவு உறுப்பமைவு உறவாகும். பின்வரும் வாக்கியத்தில் *அவன்*, *பாம்பை*, *கொன்றான்* என்ற சொற்களுக்கு இடையே உள்ள உறவு உறுப்பமைவு உறவாகும். *அவன்* என்பது எழுவாய் உறவையும் *பாம்பு* என்பது செயப்படுபொருள் உறவையும் *கொன்றான்* என்பது பயனிலை உறவையும் வெளிப்படுத்தி ஒன்றுக்கொன்று பொருண்மை உறவை வெளிப்படுத்தி நிற்கின்றன.

அவன்	பாம்பைக்	கொன்றான்
எழுவாய்	செயப்படுபொருள்	பயனிலை
செய்பவன்	அவதிப்படுவது	செயல்

அடுக்கு உறவு

ஒரு சொல்லுக்குப் பதிலாக அதே இடத்தில் வரும் சொற்களுக்கு இடையே உள்ள உறவு அடுக்குறவு எனப்படும். எடுத்துக்காட்டாகப் பின்வரும் வாக்கியத்தில் *பாம்பு*, *பல்லி*, *பூனை*, *எலி* என்பவைகளுக்கு இடையே உள்ள உறவு அடுக்குறவு எனப்படும்.

பாம்பைக்	
பல்லியைக்	
அவர் பூனையைக்	கொன்றான்
எலியைக்	

சொல்லுறவுகள்

லையான்ஸ் (1975) அடுக்குறவு முறையில் ஒன்றையொன்று இடம்பெயர்க்க இயலும் சொற்களுக்கிடையே உள்ள சொல்லுறவுகளைப் (lexical relations) பொருண்மை அடிப்படையில் பின்வரும் வகைப்படுத்தி ஆய்கிறார்.

- ஒருபொருள் பன்மொழியம் (Synonymy)
- உள்ளடங்கு மொழியம் (Hyponymy)
- இணக்கம் (Compatibility)

- இணக்கமின்மை (Incompatibility)

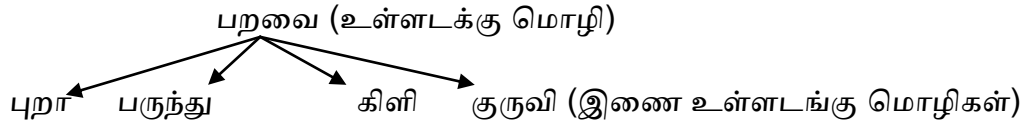
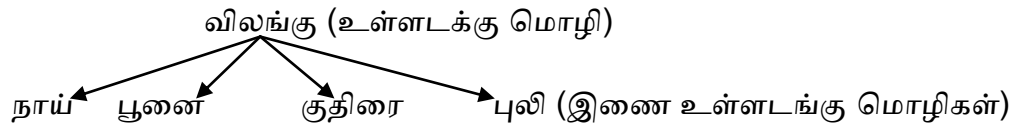
ஒருபொருள் பன்மொழியம்

பொருண்மை அடிப்படை ஒன்றையொன்று உள்ளடக்கும், இடம் பெயர்க்கும் சொற்களுக்கு இடையே உள்ள உறவு ஒருபொருள் பன்மொழியம் எனப்படும்.

புத்தகம்:நூல், உதிர்:விழு, காண்:பார், நுழை:பிரவேசி, விரட்டு: துரத்து.

உள்ளடங்கு மொழியம்

பொருண்மை அடிப்படையில் உள்ளடங்கும் சொற்களுக்கும் உள்ளடக்கும் சொற்களுக்கும் இடையே உள்ள உறவு உள்ளடங்கு மொழியம் எனப்படும்.



உள்ளடங்கு மொழியத்தை ஒரு கடவு (transitive) உறவாகக் கொள்ளலாம். 'அ' என்பது 'இ' யின் உள்ளடங்கு மொழியாகவும், 'இ' என்பது 'உ' வின் உள்ளடங்கு மொழியாகவும் வருமானால், 'அ' என்பது 'உ'வின் உள்ளடங்கு மொழியாகும்.

எ.கா.

பசுபாலூட்டி, பாலூட்டிவிலங்கு என்றால் பசுவிலங்கு ஆகும்.

இணக்கம்

பொருண்மை அடிப்படையில் ஒன்றையொன்று மேலுறலாக வரும் சொற்களுக்கு இடையே உள்ள உறவு இணக்கம் எனப்படும். இணங்கிவரும் சொற்கள் சில பொருண்மைப் பண்புகளில் ஒன்றுபட்டும் சில பொருண்மைப் பண்புகளில் வேறுபட்டும் இருக்கும்.

பாம்பு: விஷஜந்து (சரியான இணக்கம் (strict))

நாய்: செல்லப் பிராணி (தற்செயலான இணக்கம் (contigent))

இணக்கமின்மை

பொருண்மை அடிப்படையில் உள்ளடங்கவோ உள்ளடக்கவோ செய்யாமல் இணக்கமின்றி வேறுபட்டு வரும் சொற்களுக்கு இடையே உள்ள உறவு இணக்கமின்மை (incompatibility) எனப்படும்.

மலை : மடு
ஆடுதல் : தின்னுதல்

எதிர்மொழியம்

இணக்கமற்ற ஆனால் ஒன்றையொன்று எதிர்ப்பதுபோல் பொருள் அமையும் சொற்களுக்கு இடையே உள்ள உறவு எதிர்மொழியம் (antonymy) எனப்படும்.

எ.கா

உயரம் : குள்ளம்
நல்லது : கெட்டது
நீளம் : குட்டை
விரைவாக : மெதுவாக
இலகு : கடினம்

எதிர்மொழியத்தின் உட்பிரிவுகள்

எதிர்மொழியங்களை மூன்றாகப் பிரிக்கலாம்.

1. துருவ எதிர்மொழியம் (polar antonyms)
2. மேலுறல் எதிர்மொழியம் (overlapping antonyms)
3. சமசக்தியுள்ள எதிர்மொழியம் (equipollent antonyms)

துருவ எதிர்மொழியம் (polar antonyms)

எடுத்துக்காட்டாக, பின்வரும் இணைகள் துருவ எதிர்மொழியத்தின் கண் படும். கனம்:இலேசு, விரைவாக:மெதுவாக, உயரம்:குள்ளம், அகலம்:குறுக்கம், கட்டி:மெலிது, சிரமம்:இலகு

துருவ எதிர்மொழியம் பின்வருமாறு வரும்.

அவன் குள்ளமானவன் ஆனால் அவளைவிட உயரமானவன்.

அவன் உயரமானவன், ஆனால் அவளைவிட குள்ளமானவன்.

மேலுறல் எதிர்மொழியம் (overlapping antonyms)

மண்டு:புத்திசாலி, நல்லது:கெட்டது, இரக்கம்:கொடுரம்

அவன் மண்டு, ஆனால் அவளைவிட புத்திசாலி

அவன் புத்திசாலி, ஆனால் அவளைவிட மண்டு

துணைநிலை (complementarily)

துணைநிலைச் சொற்கள் ஒரு கருத்துப்பரப்பை (conceptual domain) ஒன்றுக்கொன்று தனிப்பட்ட இரண்டு பகுதிகளாகப் பிரிக்கும்.

உண்மை:பொய், திற:மூடு, வெற்றி:தோல்வி

இது இணை துணைச் சொற்களில் ஒன்று காட்டும் பொருளை ஏற்றால் மற்றதன் பொருளை மறுப்பதன் ஒக்கும்.

கதவு திறந்திருக்கிறது என்றால் கதவு மூடியிருக்கவில்லை என்று அர்த்தம். கதவு திறக்கவுமில்லை, அடைக்கவுமில்லை என்பது சரியில்லை. அவன் தேர்வில் வெற்றியடைந்தான் என்றால் அவன் தேர்வில் தோல்வியடையவில்லை என்று அர்த்தம். அவன் தேர்வில் வெற்றியடைவுமில்லை தோல்வியடையவுமில்லை என்பது சரியல்ல.

மறுதலை (conversances)

இரண்டு சொற்களில் ஒன்றின் பொருள் மற்றதின் மறுதலையாக வருமானால் அவைகளுக்கு இடையே உள்ள உறவை மறுதலை என்கிறோம்.

கணவன்:மனைவி, குழந்தை:பெற்றோர்

இராஜா அவள் கணவன் என்றால் இராஜாவின் மனைவி என்று அர்த்தம். இராஜா அவர்கள் குழந்தை என்றால் அவர்கள் இராஜாவின் பெற்றோர்கள் என்று அர்த்தம்.

பரஸ்பர சமூக பாத்திரங்கள்

வைத்தியர்:நோயாளி, முதலாளி:தொழிலாளி, எஜமானி:வேலைக்காரி

உறவுமுறைகள்

அப்பா/அம்மா: மகன்/மகள்

கால இடைநிலை உறவுகள்

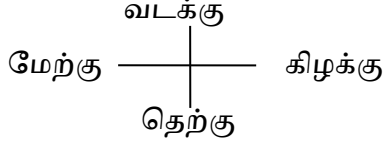
முன்னே:பின்னே, முன்னால்:பின்னால், மேலே:கீழே

திசை எதிர்மறைகள் (Directional oppositions)

வா	:	போ
மேலேறு	:	கீழிறங்கு
முன்னேறு	:	பின்னேறு
சேர்	:	புறப்படு
வலது	:	இடது
முன்	:	பின்

செங்கோண எதிர்நிலை உறவும் துருவ எதிர்நிலை உறவும் (orthogonal and antipodal opposition)

இவ்வகையில் படும் எதிர்நிலைகளைத் திசைச் சொற்களால் விளக்கலாம்.



செங்கோண எதிர்நிலை உறவு: இவ்வறவு காட்டும் இணைகளில் ஒவ்வொரு சொல்லும் பிற இரண்டு சொற்களுடன் செங்கோண எதிர்நிலையில் வரும்.

வடக்கு:கிழக்கு & மேற்கு, கிழக்கு:தெற்கு & வடக்கு

சிறுமி:சிறுவன் & பெண், girl: boy & wowan

துருவ எதிர்நிலை உறவு: இவ்வறவு காட்டும் இணைகளில் ஒவ்வொரு சொல்லும் மற்ற சொல்லுடன் துருவ எதிர்நிலையில் உள்ளது.

வடக்கு:தெற்கு, கிழக்கு:மேற்கு

மேலே:கீழே, முன்னால்:பின்னால், இடது:வலது

துருவ எதிர்சொற்கள் இடச்சொற்களை மட்டுமின்றி பிற சொற்களையும் உள்ளடக்குகிறது.

எ.கா.

நிறங்கள் – கறுப்பு:வெள்ளை, சிவப்பு: கறுப்பு

எதிர்மொழி என்பது 'ஒருபோல்' என்பதிலிருந்து மிக வேறப்பட்டு எதிர்ப்புறமாக இருப்பது என்று கருதப்பட்டது. அது சரியல்ல. நாம் இரண்டு பொருள்களை ஒப்பிடவோ வேறுபடுத்தவோ செய்கையில் நாம் அவைகளுக்குள்ள வேறுபாட்டையும் ஒற்றுமையும் பார்த்து என்ன குணங்களில் வேறுபடுகின்றன என்று கூறுகிறோம்.

எ.கா.

married : single

ஒற்றுமைகளை நோக்கித்தான் எதிர்மொழிய உறவு நிறுவப்படுகிறது.

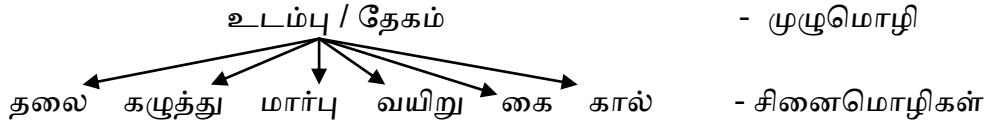
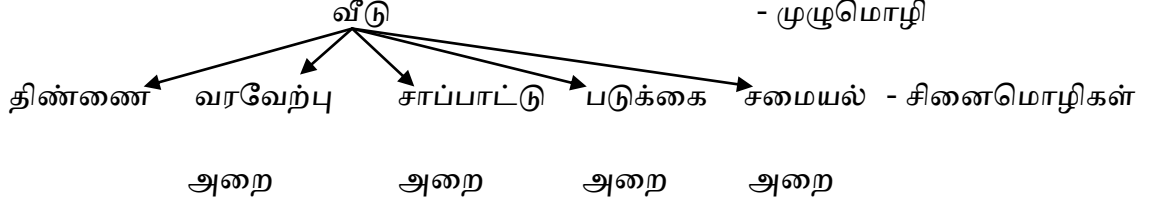
பகுதி-முழுமை உறவுகள் (part and whole relations)

பகுதி-முழுமை உள்ளடக்கு-உள்ளடங்கு உறவிலிருந்து மாறுபட்டாகும். எடுத்துக்காட்டாகப் பின்வரும் இணைகள் பகுதி-முழுமை உறவை வெளிப்படுத்தும்:

கை:உடம்பு, சக்கரம்:வண்டி.

கை ஒருவகை உடம்பு அல்ல. உடம்பின் உறுப்பு தான். உடம்பின் ஒரு உறுப்பு கை ஆகும். பகுதி-முழுமை உறவு உள்ளடங்கு உறவைப் போல கடவு (transitive) உறவாகும். 'அ' என்பது 'ஆ' என்பதன் பாகம் என்றும் 'ஆ' என்பது 'இ' வின் பாகம் என்று கொண்டால் 'அ' என்பது 'இ' என்பதன் பாகமாகும். எடுத்துக்காட்டாக, கண் என்பது தலை என்பதன் உறுப்பு என்றும் தலை என்பது உடம்பு என்பதன் உறுப்பு

என்றும் கொண்டால் கண் என்பது உடம்பு என்பதன் உறுப்பாகும். உறுப்பை அல்லது சினையக் குறிப்பிடும் சொற்களை உறுப்பு மொழிகள் அல்லது சினையமொழிகள் என்றும் முமையைக் குறிப்பிடும் சொற்களை முழுமையமொழிகள் என்றும் குறிப்பிட இயலும். பல முழுப்பெயர்களை உடம்பு போல் பல உறுப்புகளாகப் பிரிக்கலாம். எடுத்துக்காட்டாக வீடும் வீட்டின் பாகங்களையும் கூறலாம்.



பகுதி-முழுமை போன்ற உறவுகள்

பகுதி-முழுமை போன்ற பல்வேறு உறவுகள் சொற்களுக்கிடையில் வருவதைக் காணலாம். இவ்வேறுபாடுகளைப் பல பரிணாமங்களில் உணரலாம்.

1. இடம் – இடத்தின் ஒருபகுதி

தாலுகா: மாவட்டம்: இராஜியம்: தேசம்

தலைநகர்: தேசம்

2. குழு – அங்கத்தினர் உறவு (group-member relation)

குடும்பம்: குடும்பத்தினர்

3. வகுப்பு – அங்கத்தினர் உறவு (class-member relation)

குவியல்: நெல்

காடு: மரம்

நூலகம்: நூல்

4. உறுப்புகள் அல்லது கூட்டுப்பொருளின் ஒரு பகுதி (constituents or ingredients)

மசாலா: மிளகு, ஜீரகம்

5. பொருள் : மூலப்பொருள் (object : material)

ஆடை: நூல்

6. பொருள் : துகள் (substance : particle)

மழை: துளி

மேற் கூறிய சொற்களுக்கிடையிலே உள்ள பொருண்மை உறவுகள் அடிப்படையில் சொற்களஞ்சியம் அமைப்பாக்கம் செய்யப்பட்டுள்ளது.

2.3.நைடாவின் பொருண்மைக் கோட்பாடு

நைடா (1975a) குறிப்புப்பொருளின் (Referential meaning) பொருளணுப் பகுப்பாய்வினை ஒரு மொழியின் சொற்றொகையின் பொருண்மை அமைப்பைக் காண்பதற்குப் பயன்படுத்துகிறார். சொல்லுக்கும் அதன் குறிப்புக்கும் (referent) உள்ள தொடர்புதான் குறிப்புப் பொருளாகும். குறிப்புப் பொருள் இரண்டு காரணிகளை அடிப்படையாகக் கொண்டது. அவை (உணர்ச்சிப் பொருளுக்கு எதிரான) புலனறிவுப் பொருளும் (இலக்கணப்பொருளுக்கு எதிரான) குறிப்புப் பொருளும் ஆகும். குறிப்புப் பொருள் ஒரு சொல்லின் கட்டமைந்த ஒரு கூட்டம் புலன்கூறுகள் (cognitive features) அடங்கியதாகும். இப்புலன்கூறுகள் ஒரு சொல் மற்றொரு சொல்லில் இருந்து வேறுபடப் பயன்படுகின்றன. ஒரு பொருள் பரப்பில் (semantic domain) அடங்கும் சொற்கள் தங்களுக்குள் வேறுபட உதவும் இம்முக்கியமான பொருண்மைக் கூறுகளை நாம் கண்டுபிடிக்கவேண்டும். ஒத்த பொருண்மைக் கூறுகளைக் கொண்டிருக்கும் குறிப்புப் பொருள்களை ஒன்றாக ஒரு பொருள் பரப்பில் அடக்கலாம். இப்பொருள் பரப்பின் தேவை (Relevanc), பருப்பம் (largeness), படிநிலை அமைப்பில் அதன் செயற்பாட்டு நிலை மட்டம் (level of function) ஆகியன மொழியின் மொத்த பொருண்மைக் கட்டமைப்பைப் பொறுத்திருக்கின்றன. ஒன்றுக்கொன்று தொடர்புடைய பொருள் தொகைகளின் பொருண்மைத் தொடர்புகளின் விளக்கம் இரண்டு வேறு பட்ட வகையான உறவுகளை எடுத்துக்கொள்ளவேண்டும். அவையாவன வெவ்வேறு சொற்களுக்கு இடையே உள்ள தொடர்புடைய பொருள்கள் மற்றும் ஒரு சொல்லின் தொடர்புடைய பொருள்களுக்கு இடையே உள்ள உறவு. ஒரு பொருள் பரப்பில் அடங்கும் வேறுபட்ட சொற்களின் பொருள்களை நான்கு வழிகளில் உறவுபடுத்தலாம். அவை உள்ளடங்கல் (inclusion), மேலுறல் (overlapping), துணைநிலைப்படுத்தல் (complementation) அண்மைப்படல் (contiguity) ஆகியன. உள்ளடங்கல் படிநிலைத் தொடரில் தங்களுக்குள் தொடர்புள்ள மற்றும் அவை அடங்கியுள்ள பொருளிலிருந்து குறைந்தது ஒரு வேறுபடுத்தும் பொருளணு கூடுதலாக உள்ள பொருள்கள் அடங்கிய தொகுதியை உருவாக்கும் (எ.கா. விலங்கு:பாலூட்டி). மேலுறல் ஒரு படிநிலை மட்டத்தில், மரபுப்படி ஒருபொருள் பன்மொழிகள் என்று அறியப்படும் மற்றும் அவற்றின் மேலுறல் தொகுதியை உருவாக்கும் (எ.கா.

கொடு:அளி, காய்ச்சல்:சுகக்கேடு). துணைநிலை பொருள்களின் தொகுதிகள் மூன்று முக்கிய வழிகளில் தொடர்புடையன. அவை எதிர்நிலைச் சொற்கள் (opposites) (எ.கா. பெரிது:சிறிது), திருப்புப்பொருள் நிலைச் சொற்கள் (reversies) (எ.கா. கட்டு:அவிழ்), மறுதலைச்சொற்கள் (conversies) (எ.கா. வாங்குதல்: விற்றல்) ஆகியன. ஒரு படிமட்டத்தில் கூடுதல் பொருள் ஒப்புமை உள்ள அதேசமயம் குறைந்தது ஒரு முக்கிய பொருளணுவால் குறிப்பிடத்தக்க விதத்தில் வேறுபடும் பொருள்கள் ஒரு அண்மைப்பட்ட தொகுதியை உருவாக்கும் (எ.கா. நாற்காலி: பெஞ்சு:ஸ்டூல்:சோபா).

ஒரு சொல்லின் வேறுபட்ட பொருள்களை நான்கு வழிகளில் தொடர்புபடுத்தலாம்: ஆக்கம் (derivation), மாற்றீடு/இடப்பெயர்ச்சி (replacement), உருவகநீட்சி (figurative extension) மற்றும் புறஎல்லைக்கோவை (peripheral clustering). அடிப்படைப் பொருளின் பொருளணுக்களை அதனின்றும் பெறப்பட்ட வேறு பொருள் பரப்பைச் சார்ந்த பொருளில் அடக்குவது ஆக்கமாகும் (எ.கா. காதல்: காதலன்). அடிப்படைப் பொருளின் குறைந்தது ஒரு பொருளணுவால் பொருள் பரப்பில் பெரிய மாற்றம் விளைவிக்காமல் இடப்பெயர்ச்சி செய்வது மாற்றீடாகும் (எ.கா. ராஜா:ராணி). பெரிதும் வேறுபட்ட பொருள் பரப்புகளில் அடங்கும் அடிப்படைப் பொருளுக்கும் உருவகப்பொருளுக்கும் உள்ள தொடர்பைச் சில துணைப் பொருளணுக்கள் (supplementary components) அல்லது திருப்பிப் பொருள் கொள்ளப்பட்ட பொருளணுக்கள் (reinterpreted diagnostic components) வாயிலாக நிறுவுவது பொருளின் உருவக நீட்சியாகும் (எ.கா. நரி:தந்திரசாலி). மையப் பொருளுடன் (central meaning) தொடர்புடைய அல்லது அத்துடன் சிலப் பொதுப் பொருளணுக்களைக் கொண்டு தங்களுக்குள் தொடர்புடைய புற எல்லைப் பொருள்களைக் கோவைப்படுத்துதல் புற எல்லைக் கோவையாகும் (எ.கா. கெட்டியான தோல், கெட்டியான பால் இவற்றில் வரும் கெட்டி).

நைடா (1975b:152) பின்வருமாறு கூறுகிறார்: பொருள் உணர்த்தும் பலவகைச் சொற்சேர்க்கையின் ஆய்வு நம்மை நான்கு அடிப்படை பொருண்மைக் குறியீடுகளின் வகுப்புகளை இனங்கண்டுகொள்ளத் தூண்டும். இவ்வகுப்புகளில் அடங்குபவை: 1. குதிரை, மரம், இலை, சூரியன், மலை போன்ற 'பருப்பொருள்கள் (objects)', 2. ஆடுதல், நடத்தல், வெட்டுதல், தள்ளுதல், போதல், இருத்தல், புரள்தல், போன்ற 'நிகழ்வுகள் (events)', 3. அ) சிவப்பு, நீலம், சின்னது, பல, ஒன்று, இரண்டு (பருப்பொருள்களின் அருவங்கள்), ஆ) விரைவு, மெல்ல, ஒருதடவை, இரண்டுதடவை, அடிக்கடி (நிகழ்வுகளின் அருவங்கள்), இ) கடினம் (எ.கா. கடினமான பாறை, கடினமான வேலை),

மிருது (எ.கா. மிருதுவான கோட்டு, மிருதுவான வேலை) (பருப்பொருள்கள் அல்லது நிகழ்வுகள் என்பதன் அருவங்கள்) அல்லது ஈ) மிக (எ.கா.) மிகப்பல) ரெம்ப (எ.கா. ரெம்ப சின்ன), வெகு (எ.கா. வெகு விரைவு) (அருவங்களின் அருவங்கள்) இவற்றைத் தொடர்புபடுத்த உதவும் 'தொடர்பன்கள்' (events)', எ.கா. அ) பருப்பொருளுடன் பருப்பொருள்: ஜானும் பில்லும், வீட்டிலுள்ள மனிதர்கள், மேஜை மேலுள்ள புத்தகம்; ஆ) பருப்பொருளுடன் நிகழ்வு: பறக்கும் மனிதர்கள், வேட்டை நாய்களின் குரைப்பு; இ) நிகழ்வுடன் நிகழ்வு: ஓட்டமும் குதிரையும், நீச்சலில் வெற்றி; ஈ) பருப்பொருளுடன் அருவம்: பொருளின் அழகு, ரோஜாவின் நிறம்; உ) அருவத்துடன் அருவம்: மோசமும் சோம்பலும், பரிசுத்தஅழகு. இந்த நான்கு அடிப்படைப் பொருண்மை வகுப்புகளும் எல்லா மொழிகளிலும் காணப்படுவதுடன் எல்லா வகையான சொல்லுருக்களுக்கு இடையே உள்ள பொருள் தொடர்பை ஆய்வதற்கும் அடிப்படையாக அமையும் என்றும் தோன்றுகின்றது.”

நைடா கிரேக்கப் புதிய வேதாகமத்திற்கு ஒரு புதிய அகராதி உருவாக்க முனைந்து அதை நிறைவேற்றியுள்ளார். இது பொருள்களைச் சொல் அலகுகளின் அகரவரிசை நிரக்கில் அல்லாமல் பொருள் பரப்பின் அடிப்படையில் முறைப்படுத்தும். வெவ்வேறு பொருண்மை அலகுகளின் தொடர்பான அர்த்தங்களின் அடிப்படையில் பொருள் பரப்புகள் பகுக்கப்பட்டுள்ளன. தேவையான பொருளணுக்களைக் காணுவதில் ஒரு பொருள் பரப்புகள் பகுக்கப்பட்டுள்ளன. தேவையான பொருள்களைக் காணுவதில் ஒரு பொருள் பரப்பின் குறிப்பு நுட்பம் பொதுவாக எந்த ஒருபொருள் தொகுதியும் தம்முள் அடக்கியுள்ள பொதுவான பொருளணுக்களின் எண்ணிக்கையோடு தகவுப்பொருத்தம் கொண்டிருக்கும். இந்த அகராதி முக்கியமாக மொழிகளுக்கு இடையே உள்ள பொருளணு அடிப்படையிலான வேறுபாடுகளை அறிந்த பைபிள் மொழிபெயர்ப்பாளர்களுக்காகத் திட்டமிடப்பட்டதாகும். நாம் முன்னர் கண்ட சொல் உறவுகளின் அடிப்படையில் பொருளணு ஆய்வைப் பயன்படுத்தி அவர் கிரேக்க மொழியின் சொற் கட்டமைப்பைத் தர முயன்றுள்ளார். கீழே தரப்பட்டுள்ளது நைடா (1975a:176-186) தந்துள்ள வகைப்பாட்டின் சுருக்கமாகும்.

நைடா (1975a) அண்மைப்படுதல் (contiguity), மேலாறல் (overlapping), உள்ளடங்கல் (inclusion), துணைநிலைப்படுதல் (complementation) ஆகிய உறவுகளால் சொற்றொகையைப் பாகுபடுத்துகிறார். நைடா அண்மைப்படுதலை அடிப்படை உறவாகக் கருதுகிறார்.

நைடா குறிப்புப் பொருளின் (referential meaning) வேறுபடல் பொருள் கூறு ஆய்வை (componential analysis of meaning) ஒரு மொழியின் சொற்றொகையின் பொருண்மை அமைப்பைக் காண்பதற்குப் பயன்படுத்துகிறார். சொல்லுக்கும் அதன் குறிப்புக்கும் உள்ள தொடர்பு தான் குறிப்புப் பொருளாகும்.

நைடா (1975:152) பொருள் உணர்த்தும் பலவகைச் சொற்சேர்க்கைகள் ஆய்வு நம்மை நான்கு அடிப்படைப் பொருண்மைக் குறியீடுகளின் வகுப்புகளை இனங்கண்டு கொள்ளத் தூண்டும்.

1. பருப்பொருள்கள்

எ.கா. குதிரை, மரம், இலை.

2. நிகழ்வுகள் (வினைகள்)

எ.கா. ஓடுதல், நடத்தல், வெட்டுதல்.

3. பெயரடைகள் (பருப்பொருள்களின் அருவங்கள்)

எ.கா. சிவப்பு, நீலம், சின்னது.

4. வினையடைகள் (நிகழ்வுகளின் அருவங்கள்)

எ.கா. விரைவு, மெல்ல.

5. பல பருப்பொருள், நிகழ்வுகள், அருவங்கள் இவற்றைத் தொடர்புபடுத்த உதவும் தொடர்பான்கள்.

எ.கா வீட்டிலுள்ள மனிதர்கள்

பறக்கும் மனிதர்கள்

ஓட்டமும் குதியும்

பொருளின் அழகு

இந்த நான்கு பொருண்மை அமைப்புகளும் எல்லா மொழிகளிலும் காணப்படுகிறது.

இவை எல்லா வகையான சொல்லுருக்களுக்கிடையேயுள்ள பொருள் அடிப்படையிலான தொடர்பை ஆய்வதற்கு அடிப்படையாக அமைகிறது.

நைடாவின் பொருண்மையமைப்பு/பொருட்புலப் பாகுபாடு

1. பருப்பொருள்

எ. விலங்குகளல்லாதவை

1. இயற்கையானவை

எ. பூகோளம் சம்பந்தப்பட்டவை

பி. இயற்கைச் சாதனங்கள்

சி. தாவரங்கள் மற்றும் தாவரப்பொருட்கள்

2. உருவாக்கிய அல்லது கட்டப்பட்ட பொருட்கள்

எ. கலைப் பொருட்கள்

பி. பதன் செய்யப்பட்ட பொருட்கள்: உணவுகள் மற்றும்

வாசனைப்பொருட்கள்

சி. கட்டிடங்கள்

பி. விலங்குப்பொருட்கள்

1. விலங்குகள், பறவைகள், பூச்சிகள்

2. மனிதர்கள்

3. இயற்கைக்கு அப்பாற்பட்ட சக்திகள் அல்லது உயிர்கள்

2. நிகழ்வுகள்

பௌதிகம் தொடர்பானவை, உடற்கூறு தொடர்பானவை, உணர்வு தொடர்பானவை, உணர்ச்சி தொடர்பானவை, அறிவு தொடர்பானவை, கருத்துப்பரிமாற்றம், சோக்கை, கட்டுப்பாடு, சலனம், தாக்கம், கைமாற்றம், தொடர்ச்சியான சலனங்கள் அல்லது செயல்கள் உள்ளடங்கிய பல்கூட்டச்செயல்கள்

3. அருவங்கள்

காலம், தூரம், பருமை, விரைவு, சூடு, நிறம், எண், நிலை, மதப்பண்பு, கவர்ச்சி, வயது, உண்மை-பொய், நல்லது-கெட்டது, பலம், ஆரோக்கிய நிலை.

4. தொடர்புபடிகள்

இடம் தொடர்பானவை, காலம் தொடர்பானவை, சுட்டல், தர்க்கம் தொடர்பானவை.

2.4 சொற்களஞ்சியம் உருவாக்கம்

சொற்களஞ்சியம் என்று அழைக்கப்படுகின்ற பொருட்புல அகராதி அதன் விரிந்த தற்காலப் பொருளில் சொற்களைக் கருத்துக்கள் தலைப்புகள் அல்லது பாடப்பொருட்கள் அடிப்படையில் பாகுபாடு செய்வதாகும். சொற்களஞ்சியம் உருவாக்குகையில் கருத்தில் கொள்ள வேண்டிய சில மொழியியல் வாதப் பொருட்களும் பின்பற்ற வேண்டிய சில கணிப்பொறி வழிமுறைகளும் உள்ளன.

பின்வரும் கருத்துருக்கள் சொற்களஞ்சிய உருவாக்கத்திற்கு அடிப்படையாகும்.

1. சொற்களுக்கிடையில் உள்ள சொல் உறவுகளை நிறுவுதல்.

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

2. குறைந்த பட்சம் ஒத்தல், உள்ளடங்கல், இணக்கம், இணக்கமின்மை என்ற நான்கு சொல் உறவுகளால் ஆன படிநிலை வகைப்பாட்டை நிறுவுதல் மற்றும் வேறுபடுத்துகிற பொருண்மைப் பண்புக் கூறுகளைக் கொண்டு சொற்களை நிறுவுதல்.
3. சொற்களைத் தேர்ந்தெடுத்தல்.
4. சொற்களை இறுதி சொற்களத்தில் வரிசைப்படுத்தல்.
5. அடிப்படைப் பொருளை ஆக்கப்பொருளுடன் தொடர்புபடுத்தல்.

பொருண்மை பாகுபாட்டின் சிக்கல்கள்

நைடாவின் பொருண்மையமைப்பு/ பொருண்மை பாகுபாடு அடிப்படையில் தமிழ்ச் சொற்களைப் பாகுபாடு செய்வது சில இடங்களில் கடினமானதாக அமைகிறது. ஒரு குறிப்பிட்ட சொல் எந்தப் பொருள் பரப்பைச் சார்ந்தது என்பதை முடிவு செய்வதில் அடிக்கடி தெளிவின்மை ஏற்படுகிறது. நைடா சிக்கல்களைப் பின்வருமாறு கூறுகிறார்.

- சொல்லாக்கப் பொருளுக்கும் அதன் அடிப்படைக்கும் உள்ள உறவு
- சொல் திரிபு வேற்றுமைகளைக் கையாளுதல்
- ஒரு சொல்லின் வேறுபட்ட பொருள்கள் வேறுபட்ட மட்டங்களில் வருதல்
- சொற்பரப்புகளுக்குப் பொதுப்பொருள் வருதல் அல்லது வராமை
- படிநிலைப் பாகுபாட்டின் கட்டமைப்பு
- சொற்பரப்புகளுக்கு இடையேயுள்ள எல்லைகள்

சொற்றொகையின் அமைப்பாக்கம்

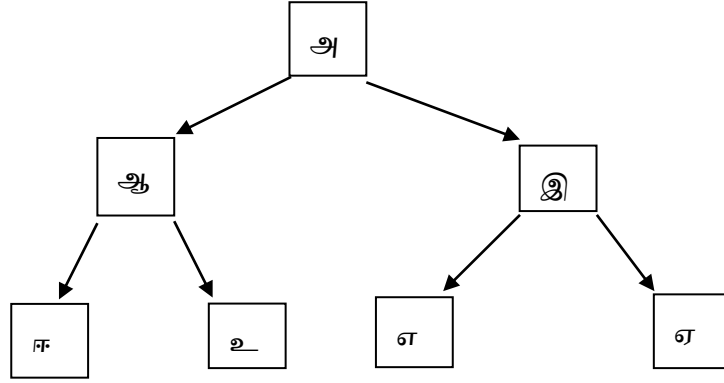
ஒரு மொழியின் சொற்றொகை (vocabulary) மனவெளியில் அங்கொன்றுமிங்கொன்றுமாக முழுவதும் சிதறிக்கிடக்கும் சொற்களின் சேமிப்பு அல்ல. இது குறைந்தது பகுதி அமைப்பாக்கம் செய்யப்பட்டது மற்றும் பல மட்டங்களில் அமைப்பாக்கம் செய்யப்பட்டதாகும். சொற்றொகையின் அமைப்பாக்கத்தின் பல பாங்குகள் இருக்கின்றன. வெளிப்படையாக அமைப்பின் இரண்டு முக்கியமான வகைகளை வேறுபடுத்துவது நன்மைபயக்கும்: மொழியியல் அமைப்பு மற்றும் உள்மொழியியல் அமைப்பு. புலனறி மொழியியல் சார்புள்ள யாரும் இவை

சுதந்திரமானவை என்று கூற விரும்பமாட்டார்கள். சொற்களஞ்சியத்திலுள்ள மொழியியல் அமைப்புகள் மொழியில் அடிப்படையில் விரையறைவிளக்கம் செய்யப்பட்டுள்ளன; நாம் இங்கு கருத்தில் கொள்வன பொருண்மை உறவுகளின் அடிப்படையில் பொருண்மையியல் விரையறைவிளக்கம் செய்யப்பட்டுள்ளன; உள்மொழியியல் அமைப்புகள் சேர்க்கைத் தொடர்புகள், தொடக்கத் தனிப்பண்புகள் மற்றும் பேச்சுத் தவறுகளின் அமைப்பொழுங்குகள் இவற்றின் அடிப்படையில் விரையறைவிளக்கம் செய்யப்பட்டுள்ளன. ஒரு பொருண்மை அமைப்பு மொழிப் பயன்பாட்டின் அமைப்பொழுங்குகளில் ஏதாவது ஒரு வழியில் பிரதிபலிக்கப்படும் மற்றும் இந்த அர்த்தத்தில் இது 'உள்வியல் அடிப்படையில் உண்மை' ஆகும். ஆனால் ஆய்வின் குறிப்பிட்ட மற்றும் தனித்தன்மையான உள்மொழியியல் உத்திகள் இதை இயைந்த அமைப்பாக வெளிப்படுத்தாது. இங்கு எடுக்கப்பட்டுள்ள நிலைமை இவ்விரு அணுகுமுறைகளும் துணைநிலையானதாகும் என்பதாகும்.

சொற்களஞ்சியத்திலுள்ள மொழியியல் அமைப்புகளுக்கு ஒலியனியல், இலக்கணவியல் அல்லது பொருண்மையில் அடிப்படை இருக்கும். இலக்கண அமைப்பாக்கத்தின் வெளிப்படையான எடுத்துக்காட்டுகள் சொல் வகுப்புகள் (தொடரியல் பண்புகள் அடிப்படையில் சொற்களை குழுமுதல்) மற்றும் சொல் குடும்பங்கள் (பொதுவான வேரிலிருந்து ஆக்கப்பட்ட சொற்களின் குழுமங்கள்) ஆகும். நாம் இங்கு பொருண்மையியல் அடிப்படையில் விரையறை விளக்கம் செய்யப்பட்ட அமைப்புகள் மீது, குறிப்பாக அர்த்த உறவுகள் அல்லது அர்த்த உறவுகளின் குழுமங்கள் மீது கருத்து செலுத்துவோம். நாம் அடுக்குறவு அர்த்த உறவுகளில் அடிப்படையில் அமைந்தவற்றில் தொடங்குவோம்.

1) படிநிலை அமைப்புகள் (hierarchies)

சொற்களஞ்சியத்திலுள்ள மிக முக்கியமான அடுக்குறவு அமைப்பின் வகைகளில் ஒன்று கிளைப் படிநிலை (branching hierarchy); இதன் மூலமுன்மாதிரி வடிவம் கீழே தரப்பட்டுள்ளது:



ஒரு தரப்பட்ட படிநிலையின் வகையை இரு உறவுகளால் பண்பாக்கம் செய்யலாம்: ஆதிக்க உறவு (relation of dominance) மற்றும் வேறுபடு உறவு (relation of differentiation). அ-வுக்கும் ஆ-வுக்கும், அ-வுக்கும் இ-க்கும், ஆ-வுக்கும் ஈ-க்கும், ஆ-வுக்கும் உ-க்கும், இ-க்கும் எ-க்கும், இ-க்கும் ஏ-க்கும் இடையிலுள்ள உறவு ஆதிக்க உறவாகும்; இது படத்தில் கணுக்களை இணைக்கும் கோடுகளால் குறியீடு செய்யப்பட்டுள்ளது. ஆ-வுக்கும் இ-க்கும், ஈ-க்கும் உ-வுக்கும், எ-க்கும் ஏ-க்கும் இடையே உள்ள உறவு வேறுபடு உறவாகும். ஒரு நல்லுருவாக்கப்பட்ட படிநிலையில், அமைப்பு முழுவதும் ஆதிக்கத்தின் மற்றும் வேறுபாட்டின் உறவுகள் நிரந்தரமாகும்.

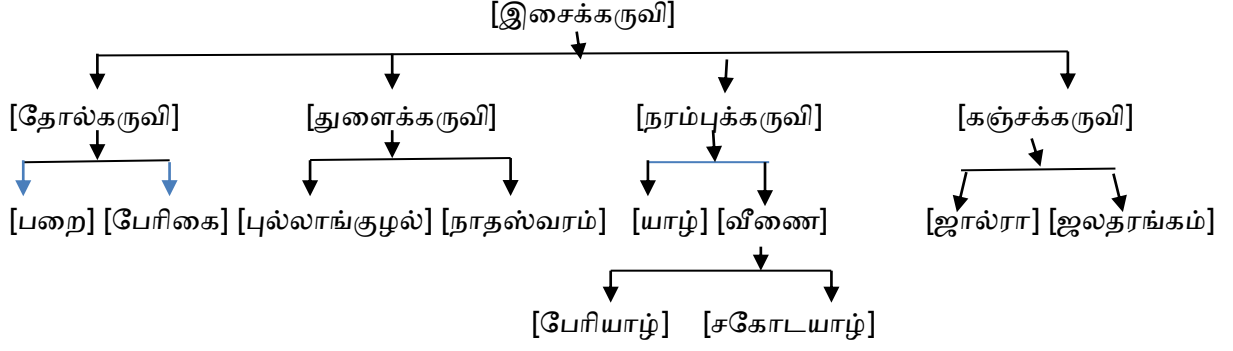
கிளைகள் அவை கீழ்நோக்கிச் செல்லுகையில் மீண்டும் சேர்வதில்லை என்பது நல்லுருவாக்கப்பட்ட படிநிலையின் (well-formed hierarchy) கூடுதலான தனிப்பண்பு ஆகும்; அதாவது [தனித்தன்மையான தாய் கட்டுப்பாடு (unique mother constraint)]

என்று அழைக்கப்படுகின்ற] மிக உயர்ந்ததைத் (தொடங்கி (beginner) என்று சில வேளைகளில் அழைக்கப்படுகின்ற அ தவிர படிநிலையிலுள்ள எந்தத் தனிமத்திற்கும் அதை உடனடியாக ஆதிக்கம் செய்யும் ஒரே ஒரு தனிமம் தான் இருக்கின்றது. உறவுகளின் குறிப்பிட்ட வகைகள் தான் இந்த நிலையை உத்திரவாதம் செய்ய இயலும்.

இங்கே கருதப்பட்ட அ, ஆ...ஏ என்பன ஒரு சொல் படிநிலை அமைப்பில் சொல் அலகுகளுடன் (அல்லது அர்த்தத்தின் அலகுடன்) பொருந்தும். சொல் படிநிலையின் இரு முக்கிய வகைகள் இருக்கின்றன: (1) ஆதிக்கத்தின் உறவு (அல்லது அதன் மறுதலை) வகைபாடாகவும் (taxonomy) வேறுபாட்டின் உறவு துணை வகைபாடாகவும் (co-taxonomy) இருக்கும் வகைப்பாட்டுப் படிநிலைகள் (taxonomic hierarchies) மற்றும் (2) ஆதிக்கத்தின் உறவு சினைமொழியமாகவும் (சினை-முழுமை) (meronymy) [அல்லது மிகச்சரியாக முழுமொழியம் (holonymy)] மற்றும் வேறுபாட்டின் உறவு துணை சினைமொழியமாக (co-meronymy) இருக்கும் சினைமொழியப் (சினை-முழுமை) படிநிலைகள்.

1) வகைப்பாட்டுப் படிநிலைகள் (taxonomic hierarchies)

வகைப்பாட்டுப் படிநிலைகள் முக்கியமாக பாகுபடுத்தும் ஒழுங்கமைப்பு; அவை ஒரு மொழியைப் பேசுபவர்கள் அனுபவத்தின் உலகத்தை வகைப்பாடு செய்வதை பிரதிபலிக்கும். ஒரு நல்லுருவாக்கப்பட்ட வகைப்பாடு சிறப்பீட்டின் வேறுபட்ட நிலைகளில் முறையான மற்றும் திறமையான வகைப்பாடுகளின் குழுமத்தைத் தருகின்றன. வகைப்பாட்டின் ஒரு (பகுதி) எடுத்துக்காட்டு படத்தில் தரப்பட்டுள்ளது.



மட்டங்கள் (Levels)

நல்லுருவாக்கப்பட்ட மட்டங்கள் இருப்பது வகைப்பாட்டுப் படிநிலைகளின் தனிப்பட்டபண்புகள் ஆகும். இவற்றை மேலே தரப்பட்டுள்ள படத்தில் தெளிவாகக் காணலாம். எடுத்துக்காட்டியுள்ளது போன்று இசைக்கருவி 1-ஆம் மட்டத்தில் இருக்கின்றது, தோல்கருவி, துளைக்கருவி போன்றன 2-ஆம் மட்டத்தில் இருக்கின்றன. பேரியாழ், சகோடயாழ் போன்றவை 3-ஆம் மட்டத்தில் இருக்கின்றன. இங்கு நான்கு மட்டங்கள் காட்டப்பட்டுள்ளன; ஆனால் இதைக் கருவிகள் என்ற பெரிய படிநிலை அமைப்பின் ஒரு பகுதி என்று வாதிடலாம். மட்டங்களை இரண்டு வழிகளில் நிறுவலாம்: மூலமுன்மாதிரியான படிநிலை (prototypical hierarchy) மற்றும் உண்மை வாழ்க்கைப் படிநிலை (real-life hierarchy). முதல் நெறிமுறைப்படி மட்டத்தைத் தீர்மானிக்க நாம் கேள்விக்குரிய தனிமத்தையும் சேர்த்து படிநிலை அமைப்பின் உயர் கணுக்களைக் கணக்கிட்டால் போதுமானது. [படிநிலையில் பிற எல்லா தனிமங்களையும் ஆதிக்கம் செய்யும் தனித்தன்மையான தனிமம் தொடங்கியாகும் (beginner).] இந்த நெறிமுறையின் படி நாம் பேரியாழ் என்பதை 4-ஆம் மட்டத்தில் நிர்ணயிக்கலாம். கணுக்களைக் கணக்கிடுவதால் நிறுவப்படும் மட்டங்களை குரூஸ் (Cruse, 1986) உத்திசார்/நுட்பம்சார் நிலைகள் (technical levels) என்கிறார். மட்டங்களின் பிற

அணுகுமுறைகள் வேறுபட்ட மட்டங்களில் தனிமங்களின் வேறுபடுத்தும் பண்புகளைப் பார்ப்பதை உட்படுத்தும். இந்த அணுகுமுறை தனித்துநிற்கும் மடங்களைத் தரும். தனிப்பட்ட பண்புகளின் பெரிய குழுமத்தை வெளிக்காட்டும் தனித்துநிற்கும் மட்டம் உளவியலாரால் அடிப்படை மட்டம் (basic level) என்றும் மனிட மொழியலாரால் பொதுபடையான மட்டம் (generic level) என்றும் அழைக்கப்படும். படிநிலையில் அடிப்படை மட்டம் மேற்கண்ட படத்தில் 3-ஆம் மட்டத்தால் எடுத்துக்காட்டப்பட்டுள்ளது.

அடிப்படை மட்டம் (Basic level)

அடிப்படை மட்டம் மற்றும் அவற்றில் வரும் தனிமங்கள் இவற்றின் முக்கியமான தனிப்பண்புகள் பின்வருவனவாகும்:

1. அடிப்படை மட்ட வகைப்பாடுகள் நல்வகைப்பாடுகளின் இரண்டு பண்களை அதிகரிக்கின்றது: உறுப்பினர்களுக்கு இடையேயுள்ள ஒற்றுமை மற்றும் சகோதரி வகைப்பாடுகளின் (sister categories) உறுப்பினர்களின் வேறுபாடுகள். உயர்மட்டத்தில் உள்ள வகைப்பாடுகளில் உள்ளொற்றுமை குறையும்; தாழ் மட்டங்களில் வெளிவேறுபாடுகள் குறையும். அடிப்படை மட்ட வகைப்பாடுகள் முழு படிநிலையிலும் திறனுடையதாகும்.
2. அடிப்படை மட்ட வகைப்பாடுகள் கற்பனை உருவத்தை உருவாக்கவியலும் மிக உயர்ந்த மட்டத்தை உருபடுத்தம் செய்யும். ஒரு யாழை வடிவமாகக் கற்பனை செய்யவியலும், ஆனால் இசைக்கருவி என்பதை (ஒரு உருப்படுத்த எடுத்துக்காட்டை தெரிந்தெடுக்காமல்) வடிவமாகக் கற்பனை செய்யவியலாது. அதுபோல ஒரு நாய்

என்பதை வடிவமாகக் கற்பனை செய்யவியலும் ஆனால் விலங்கு என்பதைக் கற்பனை செய்யவியலாது.

3. அடிப்படை மட்ட வகைப்பாடுகள் நடத்தை ஊடாட்டத்தின் தனித்தன்மையான அமைப்பொழுங்குகளை விளக்கவியலும் உயர் மட்டத்தை உருப்படுத்தம் செய்கின்றன. எவ்வாறு ஒருவர் மரச்சாமான்களின்/தட்டுமுட்டுசாமன்களின் (furniture) ஒரு தனிமத்துடன் பழகுவார் என்று அவரைப் போலச்செய்யச் சொன்னால் அவரால் அது சாத்தியமாகாது. ஆனால் நாற்காலி என்ற அடிப்படை மட்ட தனிமத்துடன் பழகுவதைப் போலச்செய்யச் சொன்னால் எந்த சிக்கலுமின்றி செய்வார். இதுபோலவே கத்தி என்பது வெட்டும் கருவி என்பதைவிடவும் குதிரை என்பது விலங்கு என்பதைவிடவும் எளிதானதாகும்.

4. அடிப்படை மட்ட தனிமங்கள் அன்றாட இயற்கையான குறிப்புரைகளில் பயன்படுத்தப்படுகின்றன. அவை இயல்பான பயன்பாட்டிற்கு வழிநிலைச் சொற்களாகும்: அடிப்படையல்லாத மட்டச் சொற்களின் பயன்பாடு குறிப்பாக உந்தப்படவேண்டும். இவ்வாறு நீ ஒரு ஆப்பிள் சாப்பிட விரும்புகிறாயா? என்பது நீ ஒரு ருசியான பழம் சாப்பிட விரும்புகிறாயா? என்று கேட்பதைவிட இயல்பானது (சில சிறப்பான சூழல்களைத் தவிர்த்து). இதுபோல யார் அந்த விலங்குக்கு இன்று உணவு தரப்போகின்றார்கள்? என்ற நடுநிலையான கேள்வியைக்காட்டிலும் யார் அந்த நாய்கு இன்று உணவு தரப்போகின்றார்கள்? என்பது பொருத்தமானதாகும்.

5. மாணிடமொழியியலர்கள் அடிப்படை மட்டத் தனிமங்கள் உருபனியல் அடிப்படையில் எளியதாகவும் சொற்றொகையின் பிற பரப்புகளின் உருவக நீட்சியாகக் கடன்வாங்கப்படாததாகவும் இருப்பதைச் சுட்டிக்காட்டுகின்றனர்.

பிறமட்டங்கள் (Other levels)

அடிப்படை மட்டத்திலிருந்து கீழே வருகின்ற மட்டங்களில் உள்ள சொற்றொகைத் தனிமங்கள் அடிப்படை மட்டத்தைவிட கூட்டுச்சொற்களாக இருப்பது கூடுதல் சாத்தியமாகும் (எடுத்துகாட்டாக *பேரியாழ்*, *சகோடயாழ்*, *மகரயாழ்*, *செங்கோட்டியாழ்* என்பனவற்றைக் கவனிக்கவும்). அடிப்படை மட்டத் தனிமங்கள் எண்ணுப்பெயர்களாக இருக்கிற படிநிலைகளில் உயர்மட்டத்திலுள்ள தனிமங்கள் திரள்பெயர்களாகப் பெரும்பாலும் இருக்கும். இது குறிப்பாக உருவாக்கப்பட்ட கலைப்பொருள்களைப் (அல்லது காணப்படுகின்ற பண்புக்கூறுகளை விட செயல்பாட்டுப் பொருண்மையைக் கொண்ட சொற்களைப்) பொறுத்தவரையில் பெரும்பாலான நேர்வாகும் (எ.கா. சமையல்பொருட்கள், தட்டுமுட்டுசாமான்கள், எழுதுபொருள்கள், உள்ளடைகள்); உயிரினங்களுக்கு இது பொருந்தாது.

மட்டங்களின் எண்ணிக்கை (Number of levels)

மானிடமொழியியலாளர்களால் செய்யப்பட்ட ஆய்வு அன்றாட மொழியல் காணப்படும் வகைப்பாட்டியல் படிநிலைகள் ஐந்து அல்லது ஆறு மட்டங்களைக்; கொண்டிருக்கின்றன எனக் காட்டுகின்றது (இந்த எண்ணிக்கையும் அசாதாரணமானது): அவை பெரும்பாலும் சிறிய பகுதிகளாக நேர்கின்றன. நாம் முன்னர் பார்த்த படிநிலை எடுத்துக்காட்டு நான்கு மட்டங்களைக் கொண்டது; நாம் கருவி என்பதை உட்டுத்தினால் ஐந்து மட்டங்களைக் கொண்டிருக்கும். எண்ணிக்கை எல்லை கலைச்சொல் சொற்றொகைக்குப் பொருந்தாது.

இடைவெளிகளும் தானியங்கு வகைபாடும் (Gap and autotaxonomy)

சொல் இடைவெளிகள் வகைபாட்டியல் படிநிலைகளில் அடிக்கடி நேர்வதாகும்; குறிப்பாக அடிப்படை மட்டங்களுக்கு மேல் நேர்வதாகும். சொல் இடைவெளி நேரும் அமைப்பின் இடத்துடன் பொருந்தும் நல்லுருவாக்கப்பட்ட கருத்துரு இருப்பதன் உள்ளுணர்வு அல்லது பிற சான்று இருந்தால் நாம் சொல் இடைவெளியைப்பற்றி பேசுவோம். எடுத்துக்காட்டாக இழை, தவழ், நட, ஓடு, துள்ளு போன்ற உள்ளடங்குமொழிகளை (hyponyms) சொற்களை உள்ளடக்கும் உள்ளடக்குமொழி (superordinate term/hypernym) தமிழில் இல்லை; அதுபோல் நீந்து, பற என்பனவற்றை உள்ளடக்குமொழி (hypernym) இல்லை. காலத்தைக் குறிப்பிடும் கருவியைக் குறிப்பிடும் அன்றாடச் சொல் இல்லை. விலங்குப் பெரும்பிரிவைக் (animal kingdom) குறிப்பிடும் அன்றாடச் சொல் இல்லை.

சில சமயங்களில் ஒரு படிநிலையில் இடைவெளி அதற்கு மேலோ கீழோ இருக்கும் தனிமத்தின் அர்த்த நீட்சியால் நிரப்பப்படும்; இவ்வாறு ஒரு தானியங்கு வகைபாட்டியல் உருவாக்கப்படும்: சொல் அலகின் ஒரு பொருண்மை மற்றொன்றின் உள்ளடக்கியாகவோ செயல்படும். பின்வருவன எடுத்துக்காட்டாகும்:

1.1. அ. நீ தீபாவளிக்கு சட்டை எடுத்தாயா?

ஆ. ஆமாம், ஒரு நல்ல பான்ட் எடுத்தேன்.

1.2. அ. நீ சட்டை மாற்றப் போகின்றாயா?

ஆ. இல்லை பான்ட் மாற்றப்போகிறேன்.

(1.1-அ-வில் சட்டை உள்ளடக்குமொழியாகப் பயன்படுத்தப்படுகின்றது.)

2.1. அ. நீ காப்பி குடித்தாயா?

ஆ. ஆமாம். தேனீர் குடித்தேன்.

2.2.அ. காப்பி குடிக்கிறாயா?

ஆ. வேண்டாம். தேனீர் குடிக்கிறேன்.

(2.1.அ-வில் காப்பி உயர்உள்ளடக்குச் சொல்லாகப் பயன்படுத்தப்படுகின்றது.)

உண்மை-வாழ்கை வகைபாடுகள் (Real-life taxonomies)

நாம் எவ்வாறு நிறைவான வகைபாட்டியல் இருக்கின்றன எனப் பார்த்தோம். இருப்பினும் உண்மை-வாழ்க்கை வகைபாட்டியல் நேரடியானது அல்ல: கிளைகள் இணைவதாகவும் பொதுச் சொல் அலகுகளின் படிநிலை தெளிவற்றதாயும் இருப்பதாகத் தோன்றுகின்றது. கிக்கலான காரணிகளில் ஒன்று முழுவதும் அல்லது பெரும்பாலும் சிறப்பீடு செய்யும் (வகை-உருவாக்கும் அதாவது வகைப்பாட்டியல்) சொற்களுடன் எல்லைக்குட்பட்ட உள்பார்வைசார்ந்த சொல்லின் இருப்பாகும். ஆடை என்ற களத்தை இந்த விஷயங்களை எடுத்துக்காட்ட பயன்படுத்தலாம். நாம் ஆடை என்பதை ஆடை வகைபாட்டியலின் தொடங்கியாக எடுக்கலாம். நாம் படிநிலை அமைப்பில் கீழ் நோக்கிச் சென்றால் பின்வருவனவற்றை அடிப்படை மட்டத்தில் நேரிடலாம்: மேல்சட்டை, நிக்கர், பவாடை, கால்சட்டை, சேலை. தோற்கருவி, துளைக்கருவி என்பனவற்றிற்கு இணையான இடைமட்ட நிலையைக் காண இயலாது. இருப்பினும் இந்த நிலை எல்லைக்குட்பட்ட பல்வேறுவகைப்பட்ட உள்பார்வைசார்ந்த சொற்களின் இருப்பால் சிக்கலாக்கப்பட்டுள்ளது.

எ.கா.

இடுப்பில் அணிபவை: சேலை, வேட்டி, பாவாடை, பைஜாமா, நிக்கர், காற்சட்டை

இடுப்புக்கு மேல் அணிபவை: சட்டை, சொக்காய், ஜம்பர்

உள்ளாடைகள்: பாவாடை, வட்டுடை, பனியன், ஜட்டி

குறிப்பிட்ட நோக்கத்திற்கு: சீருடை, மேலங்கி, விளையாட்டுடை

வெளிப்படையாகப் பார்க்கவியலும் குறிப்பிட்ட நோக்கத்திற்கு என்றல்லாத ஆடை குறித்த அன்றாடச் சொல் இல்லை; இந்த வகை ஒரு பெயரற்ற வழநிலை வகைப்பாட்டின் ஒருவகையாகச் செயல்பாடும்; அவற்றிலிருந்து விலகிச்செல்பவை மட்டும் சொன்மை அடிப்படையில் வேறுபடுத்தப்படும். இயெல்லாம் ஆடைகுறித்த சொற்களிலிருந்த நல்வடிவமாக்கப்பட்ட படிநிலை அமைப்பை உருவாக்குவதை சாத்தியமறச் செய்கின்றது. குழப்பத்தோற்றதை மட்டுப்படுத்தவேண்டுமானால் பின்வரும் விஷயங்களை நாம் மனதில் கொள்ளவேண்டும்.

1. பார்வை நிரந்தரமாக வைக்கப்பட்டால்தான் சீரான படிநிலைகள் தோன்றும்;

இது இல்லையென்றால் குறுக்குப்பாகுபாடு நேரவியலும்.

2. ஒவ்வொரு பார்வையும் ஒரு தனி படிநிலையைத் தரவியலும்.

3. வேறுபட்ட படிநிலைகள் பல வழிகளில் குறிக்கிடவியலும்.

4. குறிப்பிடப்படாத பார்வையுடன் கூடிய படிநிலைகளின் சாத்தியமான எதிர்பார்ப்புகளுடன் வகைதொகைப் படிநிலைகளில் உள்ள தனிமங்கள் முழு சொல் அர்த்தங்கள் அல்ல; ஆனால் சூழல் அடிப்படையில் சுற்றிவளைக்கப்பட்ட துணை அர்த்தங்கள்.

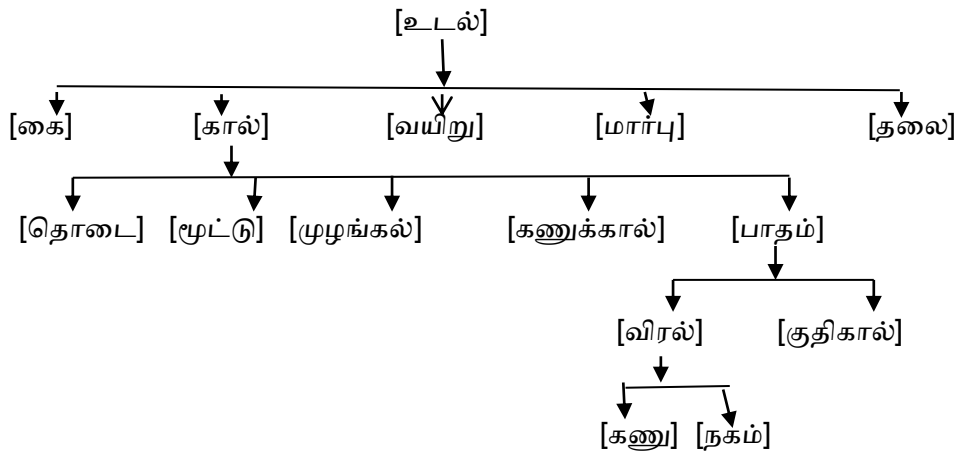
நாம் குறிப்பிட்ட பார்வைகளில் நல்லுருவாக்கப்பட்ட ஆனால் பகுதி படிநிலைகளை நிறுவுவதை எதிர்பார்க்கவியலும். எடுத்துக்காட்டாக எங்கே அணியப்படுகின்றது என்ற பார்வையில் பின்வரும் உள்ளடங்குச் சொற்களை உள்ளடக்கவியலும்:

காலணி, தலையணி, காதணி, கழுத்தணி, இடுப்பணி

இவை பொதுவான பின்னிணங்கள்/இணையும் கிளைகள் இல்லாமல் ஒன்றுக்கொன்று விலக்குவையாகும்.

2) சினைமொழியப் படிநிலைகள் (Meronymic hierarches)

சொல் படிநிலையின் இரண்டாவது முக்கிய வகை சினைமொழியமாகும் (சினை-முழு உறவாகும்); இதில் ஆதிக்கத்தின் உறவு சினைமொழியமாகும்; வேறுபடுத்தலின் உறவு துணை மொழியமாகும். விரிவான சினைமொழியங்களில் மிக அறிமுகமானது பின்வரும் படத்தில் காட்டப்பட்டுள்ள வெளிப்படையாகத் தொழும் மனித உடலின் உறுப்புகளாகும்.



இந்த படிநிலைகளின் சில விளக்கங்கள் விவாதத்திற்குரியதாகும்; எடுத்துக்காட்டாக, தோள்கள் கைகளின் பாகங்களா அல்லது உடலின் நடுப்பகுதியின் பாகங்களா என்பது. சாதாரணமாக நாம் நேரிடுகிற இயந்திரங்களுக்கு நல்லுருவாக்கப்பட்ட சினைமொழியங்கள் உள்ளன; ஆனால் வல்லுனரல்லாத மக்களில் சிலருக்குத்தான் பேருந்து, சலவை இயந்திரம் அல்லது கணிப்பொறி என்பனவற்றின் பாகங்களின் முழு

விவரத்தைத் தரவியலும். நமது அறிவின் பெரும்பான்மையும் சினைமொழியத்தின் துண்டுகள் தான்.

மட்டங்கள் (Levels)

வகைப்பாட்டியலுக்கும் சினைமொழியத்திற்கும் உள்ள முக்கியமான வேறுபாடு சினைமொழியத்தில் தெளிவான பொதுவான மட்டங்கள் இல்லாதிருப்பதாகும். கை என்பதற்கும் கால் என்பதற்கும் இடையிலுள்ள அமைப்பொப்புமை காரணமாக ஒரு அர்த்தத்தில் எடுத்துக்காட்டப்பட்ட உடல் சினைமொழியம் தனிப்பண்பற்றதாகும்: ஆங்கிலத்தில் knee என்பது elbow என்பதுடனும் sole of foot என்பது palm of hand என்பதுடனும் வடிநள என்பது க்இபநசள என்பதுடனும் பொருந்தும். (தமிழ் அடிப்பாதம்-உள்ளங்கை இணைதவிர கால் மற்றும் கைகளின் பாகங்களைக் குறிப்பிட பொதுவான சொற்களையே கொண்டள்ளது: கைமூட்டு-கால்மூட்டு, கைவிரல்-கால்விரல், அடிப்பாதம்-உள்ளங்கை, கைநகம்-கால்நகம்.) ஆனால் இது உடலின் பிற பாகங்களுக்கு நீட்சியடையவில்லை. இதன் காரணமாக ஒரு வகைபாட்டியலின் அடிப்படைக்கு நிகரன் இல்லை; சூழல் சுதந்திரமான அடையாளப்படுத்தப்படாத சிறப்பினத்தின் மட்டம் இல்லை.

சொல் இடைவெளிகள் (Lexical gaps)

வகைபாட்டியல் படிநிலையில், தொடங்கி பெரும்பாலும் சொல்நிலையாக்கம் அடையவில்லை. சினைமொழியம் இவ்வாறில்லை. இருப்பினும் இடைவெளிகள் பெரும்பாலும் தனித்தன்மையான இடத்தில் நோ;கின்றன: சிலவேளைகளில் முக்கியமான செயற்பாட்டுப் பாகம் பெயரின்றி இருக்கும். எடுத்துக்காட்டாக மூக்கு, பிடி மற்றும் மூடி இணைக்கப்படும் தேனீர்ப்பாத்திரத்தின் பாகத்தை நான் எப்படி

அழைப்போம்? சிலர் பதில் தரலாம் *பாத்திரம்* என்றோ *உடல்* என்றோ பதில் தரலாம். மற்றொரு எடுத்துக்காட்டு *கரண்டியாகும்*: ஒரு *கரண்டிக்கு* இரண்டு பாகங்கள் இருக்கின்றன: ஒன்று *பிடி* மற்றொன்று *தெரியவில்லை*.

சில யதேட்சையான இடைவெளிகள் காணப்படுகின்றன: எடுத்துக்காட்டாக விரல்கள் இணைக்கப்படும் கையின் பகுதி. இருப்பினும் இவை அரிதானவை. சில நேர்வுகளில் தானியக்க சினைமொழியத்தைக் காணலாம்; அதாவது சினையும் உடனடியான முழுமைக்கும் ஒரே பெயர் வரும். நல்ல எடுத்துக்காட்டு உடலுறுப்பு சினைமொழியத்தில் காணப்படுகின்றது: *உடம்பு* என்பது முழு உடலையும் கைகால் தலை தவிர்த்த பிற உறுப்புப் பகுதியையும் குறிப்பிடும்; *கை* என்பது முழுக்கையையும் விரல்கள் உள்ள மணிக்கட்டிற்குக் கீழேயுள்ள பகுதியையும் குறிப்பிடும்.

முரண்படும் நோக்கு (Contrastive aspects)

உள்ளுறையும் கொள்கைகள் ஏறக்குறைய உலகவயமானது எனக் கருதவியலும் என்றாலும் மொழிகள் முழுமைகள் எவ்வாறு சொன்மை அடிப்படையில் வேறுபடும் பாகங்களாகப் பிரிக்கப்பட்டுள்ளன என்பதில் வேறுபாடு காட்டுகின்றன. இதன் அர்த்தம் என்னவென்றால் வேறுபாடுகள் பின்வருவனவற்றிற்கு உட்பட்டவையாகும்: (1) ஒரே சிறு அலகுகளின் வேறுபட்ட குழுமல்கள், மற்றும் (2) எவ்வாறு துணைப்பாகுபாடுகள் செய்யப்படுகின்றன என்பதில் உள்ள வேறுபாடு. முழுவதும் பொருத்தமற்ற பாகுபாடுகள் மிக அரிதானது. (1)-இன் எடுத்துக்காட்டு ஆங்கிலத்தாலும் கிரேக்கமொழியாலும் தரப்படுகின்றது. ஆங்கிலத்தில் *hand* என்பது மணிக்கட்டைவரை குறிப்பிடும் அதற்கு மேலான பாகத்தை உட்படுத்தாது; தற்கால கிரேக்கமொழியில் *xeri* கைமூட்டுவரை குறிப்பிடும். *foot* என்பதற்கும் *podì* என்பதற்கும் இணை உறவு

இருக்கின்றது: பிந்தையது கால்மூட்டுவரை குறிப்பிடும். இவ்விரு ஒழுங்குமுறையும் மூட்டுகளை உறுப்புக்களின் எல்லைகளாகப் பயன்படுத்துவதைக் கவனத்தில்கொள்ளவும். xeri என்பதன் எந்தப் பகுதி குறிப்பிடப்படுகிறது என்பதைத் தீர்மானிப்பது சூழலுக்கு விடப்பட்டுள்ளது (மயக்கம் வருவது மிக அரிது). hand என்பதுடன் பொருந்துகின்ற xeri என்பதன் பகுதி மிக முக்கியமான பகுதியாக இருப்பதாலும் மிக அதிகமாக அடிக்கடி செயல்பாடுகளில் பங்கெடுக்கின்ற பகுதியாகவும் இருப்பதாலும் பெரும்பாலான சூழல்களில் hand என்பதை xeri என்பதுடன் தொடர்புபடுத்தி மொழிபெயர்ப்பு செய்வதால் எதுவும் இழக்கப்படுவதில்லை.

ஒரு மொழி மற்றொரு மொழியைக் காட்டிலும் நுண்ணிய வேறுபாடு தரும்போது வேறுவகையான வேறுபாடு தோன்றும். எடுத்துக்காட்டாகக் காதிற்குக் கீழ் வரும் பாகமான செவிடு என்ற தமிழ்ச் சொல்லுக்கு இணையான சொல் ஆங்கிலத்தில் இல்லை. cheek என்பது செவிடு என்பதற்கு இணையாகாது. இதை சொன்மை இடைவெளி எனக் குறிபிடவேண்டுமா என்ற கேள்வி எழலாம்.

வேறுபட்ட மொழிகளின் சினைமொழிய ஒழுங்கமைப்பு ஒப்புமையுள்ள வேறுபட்ட முழுமைகளின் பகுதிகள் பெயரிடலுக்காக குழுமப்படுவது போன்று வேறுபடுகின்றன. எடுத்துக்காட்டாக பிரஞ்சு மொழியில் கதவின் பிடி, சூட்கேசின் பிடி, பம்பின் பிடி என்பவைகளுக்கு வேறுபட்ட சொற்கள் உள்ளன. அவை பெயரிடுவதற்காக ஒரே முழுமையின் ஒன்றுபோலிருக்கற பகுதிகள் வேறுபடுவதுபோல் வேறுபடுகின்றன. எடுத்துக்காட்டாக ஆங்கிலத்தில் thumb என்பது பிற விரல்களிடமிருந்து வேறுபடுத்தப்படும். தமிழில் இம்மாரியான வேறுபாடு

செய்யப்படவில்லை. ஒவ்வொரு விரலுக்கும் தனித்தனி அடை தரப்பட்டு பெயரிடப்பட்டுள்ளன: பெருவிரல், மோதிரவிரல், ஆள்காட்டிவிரல். சுண்டுவிரல்.

பல மொழிகள் கால்களின் மற்றும் கைகளின் விரல்களைத் தொடர்பற்ற சொற்களால் குறிப்பிடுகின்றன (எ.கா. ஆங்கிலத்தில் finger, toe என வேறுபடுத்தப்பட்டுள்ளன); இருப்பினும் பல பிற மொழிகள் தொடர்புள்ள சொற்களால் குறிப்பிடுகின்றன: கைவிரல், கால்விரல்.

3.) கிடைவரிசை அமைப்புகள் (Linear Structures)

இருதுருவங்கள் (Bipoles)

கிடைவரிசை அமைப்பின் எளியவகை எதிரிடைகளின் இணையாகும்.

இருதுருவ சங்கிலிகள் (Bipolar chains)

எதிர்மைகளின் இணைகள் செயன்மை செய்யும் அளவை, பண்புகளின் வேறுபட்ட அளவுகளைக் காட்டும் சொற்களில் பலவற்றை உள்ளடக்குகின்றது. மிகக்கூடுதலான அமைப்பொழுங்கு அளவையின் ஒவ்வொரு நுனியிலும் இருக்கும் எதிர்மை இருதுருவத்தின் உட்கீடையான மீயளவுச் சொற்களாகும் (அடிப்படை எதிர்மொழி இணைகளுக்கிடையில் துருவ மாற்றம் இருக்கின்றது.):

எ.கா.

நுண்ணிய சிறிய பெரிய பிரமாண்டமான

சங்கலிகளின் கூடுதலான எடுத்துக்காட்டுகள் பின்வருவனவாகும்:

அப்பழுக்கற்ற தூய்மையான அழுக்கான

அற்புதமான மேன்மையான நல்ல கெட்ட

அழகான சாதாரணமான அசிங்கமான

ஒருதுருவச் சங்கலி (monopolar chanins)

ஒருதுருவச் சங்கலிகளில் சங்கலிகளின் முனைகளில் இருக்கும் சொற்கள் எதிர்திசைகளில் இருக்கின்றன என்பதில் அர்த்தம் இல்லை. ஒருதுருவ கிடைவரிசை அமைப்புகளின் பல வேறுபட்ட வகைகள் இருக்கின்றன. அவற்றை விளக்குவதிலும் வகைப்படுத்துவதிலும் பல வழிகள் இருக்கின்றன. பின்வரும் அமைப்பொழுங்கு குருசை (Cruse, 1986) பின்பற்றுகின்றது.

அளபுகள் (degrees)

அளபுகள் வடிவம் அல்லது தீவிரம் போன்ற தொடர்ந்து அளவை செய்யப்பட்ட பண்பின் வேறுபட்ட அளபுகளை அவற்றின் பொருண்மையின் பாகமாக உட்படுத்தும் ஆனால் உட்படுத்தலின் உறவில்லை. அவற்றின் புறவெல்லைகள் தெளிவற்றவை மற்றும் அவை உள்ளூணர்வு அடிப்படையில் அவற்றின் படித்தரபடுத்தும் தன்மையை இழக்கவில்லை. நாம் உள்ளூறும் பண்பின் மதிப்புகளைக் குறிக்கும் சொற்களிலிருந்து அவ்வாறல்லாதவற்றைக் குறிக்கும் சொற்களின் நோவுகளை வேறுபடுத்தவியலும். முந்தைய வகையின் எடுத்துக்காட்டுகள் பின்வருவன: தோல்வி வெற்றி சிறப்புநிலை . இரண்டாவது வகையின் எடுத்துக்காட்டு பின்வருவனவாகும்: பாறை, குன்று, மலை; mound hillock hill mountain.

நிலைகள் (stages)

நிலைகள் ஏதோவொன்றின் வாழ்க்கைச் சக்கரத்தின் இடங்களாகும். முன்னேற்றத்தின் கருத்துச்சாயலை உட்படுத்தும்.

எ.கா.

தொடக்கநிலை இரண்டாம்நிலை உயர்நிலை

இளங்கலைப்பட்டநிலை முதுகலைப்பட்டநிலை

குழந்தைப்பருவம் இளமைப்பருவம் வாலிப்பருவம் முதுமைப்பருவம்.

முட்டை புழு கூட்டுப்புழு வண்ணத்திப்பூச்சி

அளவைகள் (measures)

அளவைகள் பகுதி-முழுமை உறவின் அடிப்படையில் அமைந்தது; ஒவ்வொரு முழுமையும் ஒன்றுபோல் இருக்கிற பாகங்களால் பரிக்கப்பட்டிருக்கும்; அடுத்திருக்கின்ற சொல்லால் குறிப்பிடப்படும் அளவைப் பண்புகளின் மதிப்புகளுக்கிடையே வடிவியல்சார்ந்த உறவு இருக்கின்றது.

எ.கா.

வினாடி, நிமிடம், மணி, நாள், வாரம், மாதம் (போன்றன)

அங்குலம், அடி, கஜம் (போன்றன)

மைல், ஆழாக்கு, உழக்கு, படி, நாழி, குறுணி, மரக்கால்

ஸ்தானங்கள் (ranks)

ஸ்தானங்களில் உள்ளூறையும் பண்பு தொடர்ந்து வேறுபடாது; ஆனால் விட்டுவிட்டு வேறுபடும். ஒரு சொல் அதனை அடுத்துவருகின்ற சொற்களைக்காட்டிலும் கூடுதலோ குறைவோ என்றில்லை.

எ.கா.

விரிவுரையாளர், உயர்விரிவுரையாளர், இணைப்பேராசிரியர், பேராசிரியர்

கர்னல், லெப்டினட் கர்னல், மேஜர், காப்டன்

மேற்சொன் நோ;வுகளில் உள்ளுறையும் பண்பை “மூப்புநிலை” போன்று கருதவியலும் ஆனால் அவை படிப்படியாகத் தொடர்ச்சியாக மாறவில்லை: ஒரு கர்னல் மற்றொரு கர்னலைவிட கூடுதல் ஸ்தானத்தை அடையவியலாது.

அடிப்படை எண்களை (ஒன்று, இரண்டு, மூன்று, நான்கு ...) இந்தத் தலைப்பில் கருதவியலும்; இவற்றில் வேறுபடும் பண்பு “எண்ணிக்கைத்தன்மை” ஆகும். இவையும் தொடர்ச்சியாக வேறுபடவில்லை: பன்னிரண்டு எண்ணிக்கை கொண்ட எந்த வகுப்பும் பன்னிரண்டு எண்ணிக்கைகொண்ட மற்றொரு வகுப்பை எண்ணிக்கையில் விஞ்சாது. வகைப்பாட்டியல் படிநிலையின் மட்டங்களும் ஸ்தான்கள் ஆகும்:

எ.கா.

வகை, இனம், பொதுவினம், குடும்பம்

தொடர்ச்சிகள் (Sequences)

மேற்சொன்ன எல்லா நோ;வுகளிலும் வரிசையில் வரும் ஒரு தனிமம் அதை முன்தொடரும் தனிமத்தைவிட ஏதோ கூடுதல் பண்பையும் அதைத் பின்தொடரும் தனிமத்தைவிட குறைந்த பண்பையும் கொண்டிருக்கின்றது. இருப்பினும், இம்மாதிரியான நேர்வு இல்லாத நிரல்படுத்தவியலும் தனிமங்களும் இருக்கின்றன; அவை தொடர்ச்சிகள் (sequences) என அழைக்கப்படும். செவ்வாய்கிழமை திங்கள்கிழமையைவிட ஏதோ கூடுதலான ஒன்றை வைத்திருக்கின்றது என்று கூறவியலாது.

எ.கா.

திங்கள் செவ்வாய் புதன் வியாழன்...

சனவரி பெப்ரவரி மார்ச்சு ஏப்ரல்...

வசந்தகாலம் கோடைகாலம் இலையுதிர்காலம் குளிர்காலம்

காலை மதியம் மாலை இரவு

இந்த வகைப்பாடுகளை கூடுதல் அக்கறையுடன் கருதப்படவேண்டாம்: ஒரு தலைப்பின் கீழ் பல குழுமங்களைக் கருதவியலும் என்பதைக் கவனத்தில் கொள்ளவும். ஒரு திருப்தியான வகைதொகை உறவு இருக்கலாம், ஆனால் அது இதுவரை கண்டுபிடிக்கப்படவில்லை: குறுக்குப்பாடுபாடுசெய்யும் பண்புக்கூறுகள் அடிப்படையில் சிந்திக்கவியலும்.

4) சட்டகங்கள் (Grids)

சட்டங்கள் மீண்டும்நேரும் பருமைத்தன்மையான அர்த்த உறவுகளால் அல்லது பொருண்மைக் கூறுகளின் மறுநோ;வுகளால் உருவாக்கப்படுகின்றன. சட்டகத்தின் அலகு அறை ஆகும்; இது நான்கு சொல் அலகுகளால் ஆனது; அவற்றில் ஒன்றை மீதி மூன்று சொல் அலகுகளால் ஊகம்செய்யவியலும். பின்வருவன செல்களின் எடுத்துக்காட்டுகளாகும்:

1. man woman 2. hand finger

ram ewe leg toe

3. நாய் குட்டி 4. எடு திருடு

யானை கன்று கொல் கொலைசெய்

இவற்றிற்குப் பின்வருமாறு பொருண்மைக்கூறு ஆய்வுகள் தரவியலும்:

1. [X] [ஆண்] [Y] [பெண்]

[Y] [ஆண்] [Y] [பெண்]

2. [X] [X] [கூறு]

[Y] [Y] [கூறு]

3. [X] [X] [இளம்]

[Y] [Y] [இளம்]

4. [V.(1)] [V.(1)] [சட்டத்திற்குப்புறம்பானது]

[V.(2)] [V.(2)] [சட்டத்திற்குப்புறம்பானது]

பின்வருவன நல்வடிவாக்கமுற்ற அறையல்ல என்பதைக் கவனிக்கவும்: பூ ரோஜா விலங்கு பூனை ஒரு அர்த்தத்தில் இங்கே வகைதொகை உறவு மீண்டும்நேருகிறது. ஆனால் ஒரு தனிமத்தை பிற மூன்று தனிமங்களிலிருந்து ஊகிக்கவியலும் என்ற அளவீடு நிறைவேற்றப்படவில்லை. ஊகம் ஒரு திசையில் சாத்தியமாகும்: பூ ரோஜா ? பூனை ஆனால் மற்றோரு திசையில் சாத்தியமல்ல:

பூ ரோஜா

விலங்கு ?

குரூஸ் (Crues, 1986) finger மற்றும் toe என்ற சொற்களை ஒப்புவடிவங்கள் (analogues) என்கிறார். பின்வருவன பிற எடுத்துக்காட்டுகளாகும்: குழுத்தலைவர் குழு அணித்தலைவர் அணி தலைமையாசிரியர் பள்ளி துணைவேந்தர் பல்கலைக்கழகம் வலதுபக்கமுள்ளவை ஒன்றுக்கொன்று ஒப்புவடிவுகளாகும். மேலே எடுத்துக்காட்டப்பட்ட சட்டங்கள் அடுக்குறவு அடிப்படையில் சீரான போக்குடையது. ஆனால் சட்டக் கருத்துச்சாயல் அடுக்குறவு சீரான போக்கை புகுத்தவில்லை. பின்வருவன நல்லுருவாக்கப்பட்ட சட்ட அறைகளாகும்:

பேனா : எழுது பறவை: பற நாய் : குரை

மண்வெட்டி: வெட்டு மீன் : நீந்து யானை; பிளிறு

இருப்பினும் ஒப்புவடிவங்களுக்கிடையில் அடுக்குறவு இருக்கவேண்டும்.

5). கொத்துகள் (clusters)

கொத்துகள் ஒருபொருள்பன்மொழிகளின் குழுக்களாகும். இதை அமைப்பாக்கம் செய்வதன் துல்லியமும் கலவைத்தன்மையும் (சிக்கலும்) களங்களின் பிற வகைகளைக்காட்டிலும் குறைவாகும் என்பதை இது காட்டுகின்றது. இரண்டு முக்கியமான கொத்துகள் இருக்கின்றன: மையப்படுத்தப்பட்ட கொத்து (centred cluster) மற்றும் மையப்படுத்தப்படாத கொத்து (non-centred clusters).

மையப்படுத்தப்பட்ட கொத்துகள் (Centred clusters)

மையப்படுத்தப்பட்ட கொத்துகள் ஒன்றோ அதற்குமேலோ சொற்களின் ஓரளவுக்குத் தெளிவான கருமையத்தைக் கொண்டிருக்கும் மற்றும் புறஎல்லை தனிமங்களின் குறைந்த கருமையத்தையும் கொண்டிருக்கும். கருமையத் தனிமங்களின் தனிப்பண்புகளில் சில பின்வருவனவாகும்:

1. அவை வெளிப்பாட்டு அடிப்படையில் நடுநிலையானவை
2. அவை நடையியல் அடிப்படையில் அடையாளப்படுத்தப்படாதவை; அதாவது அவை பிற எந்தச் சொற்களைக் காட்டிலும் அகலமான எல்லைப் பதிவுகளில் நேரும்.
3. அவை தேற்றக்கூற்று அடிப்படையில் உயர்உள்ளடக்குச் சொற்கள்.

இற, சா, மரணமடை, மாள், சா, துஞ்சு, மண்மையைப்போடு என்ற குழுமத்தில் இற என்பது மைய உறுப்பினராகும்: அது வெளிப்பாட்டு அடிப்படையில் நடுநிலையானது மற்றும் நடையியல் அடிப்படையில் அடையாளப்படுத்தப்படாதது. மூன்றாவது

பண்புக்கூறைப் பயன்படுத்தவியலாதது; ஏனென்றால் இந்தக் குழுமத்தின் உறுப்பினர்கள் யாவும் தேற்றக்கூற்று ஒருபொருள்பன்மொழிகள்.

walk, amble, stroll, stride, saunter என்ற குழுமத்தில் *stroll*, என்பது மையச்சொல்லாகும்: இந்தக் குழுமத்தில் அடையாளப்படுத்தப்பட்ட வெளிப்பாட்டு மாற்றம் இல்லை; ஆனால் *stroll*, என்பது நடையியல் அடிப்படையில் அடையாளப்படுத்தப்படாததாகும் மற்றும் பிற எல்லாச் சொற்களின் உயர்உள்ளடக்குச் சொல்லாகும்.

மையப்படுத்தப்படாத கொத்துகள் (Non-centred clusters)

மையப்படுத்தப்படாத கொத்துகளில் சொற்கள் எல்லாம் அர்த்தக் கற்றையின் மீது பரப்பப்பட்டுள்ளன; உயர்உள்ளடக்குச் சொல் இல்லை. இவை நேரடியாக மிகக்குறைந்த அளவு தேற்றக்கூற்று வேறுபாட்டை வெளிக்காட்டுகின்றன; அர்த்தக்கற்றையின் மீது அவை ஓரளவுக்கு அண்மைப்பட்டு இருப்பது வரை இது ஒருபொருள்பன்மொழியத்தைக் குறைக்காது. பின்வருவன மாதிரி எடுத்துக்காட்டுகளாகும்:

அடி, இடி, குத்து, உதை, மோது, முட்டு, கொட்டு, குட்டு... சயி, வயி,

rap, tap, knock, slap, thwack, crack, bang, thump, bump, pop, tick, click, ring,

tinkle, clink, clank, jingle, jangle, ping...

கொத்துக்கள் மேலுறல் செய்யலாம்: இது இயல்பற்றதாகும் மற்றும் வகைதொகை மற்றும் சினைமுதல் படிநிலைகளில் விதிமுறை தழுவாததாகும். எடுத்துக்காட்டாக பின்வரும் இரண்டு கொத்துகள் மேலுறல் செய்யும்:

1 . unusual, rare, uncommon, infrequent, ...

2 . odd, queer, strange, weird, peculiar, extradentary, alien, ...

1-ஆம் குழுச் சொற்கள் அவற்றின் நோ;வில் குறைந்த நிகழ்வெண்ணைக் குறிப்பிடுகின்றன. ஆனால் 2-ஆம் குழுச் சொற்கள் பழக்கின்மையைக் குறிப்பிடுகின்றன.

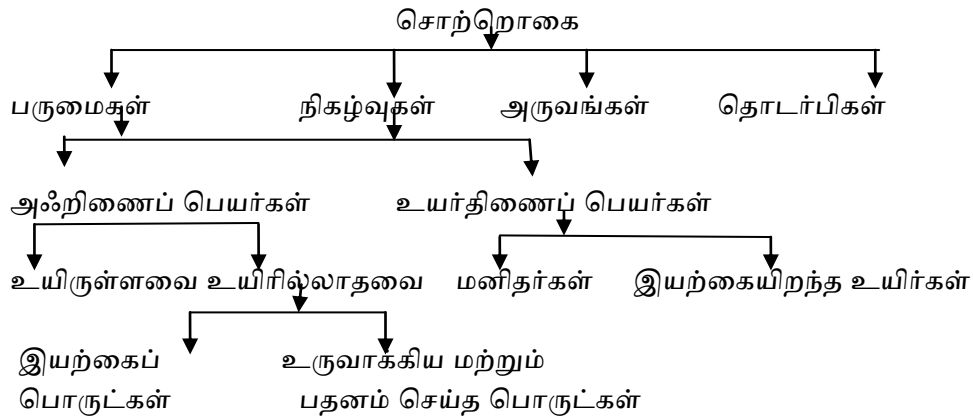
புறம்பான வகைகள் (Miscellaneous types)

நாம் சிறப்பான அமைப்புகளின் அடிப்படையில் எடுத்துக்கொள்ளவியலும் சொற்களங்களின் முக்கியமான வகைகளைப்பற்றிப் பார்த்தோம். சொற்களின் பிற முக்கியமான குழுமங்களும் உள்ளன. முதலாவது சொல் குடும்பங்கள் என்று அழைக்கப்படுபவை. இச்சொற்கள் பொதுவன ஒரு வேரிலிருந்து ஆக்கப்பட்டதாகும்: எ.கா.நடி (வி) , நடிப்பு (பெ), நடிகள் (பெ), நடிகை(பெ). இங்கு பிற வோ;களுடன் மீண்டும் நிகழும் பொருண்மையியல் செயன்மைகள் இருக்கின்றன; இந்த சொற்களை ஒரு குழும சொற்கள் என்று மட்டும் தான் கூறவியலும். இரண்டாவது, பேச்சு அல்லது முறையான சூழலில் பயன்படுத்தப்படுகின்ற வரம்பு வழக்கு (சநபளைவநச) அல்லது ஒரு மதச்சார்பான பிரச்சாரம், சட்ட ஆணவம் அல்லது மருத்துவ பாடநூல் இவற்றிற்குப் பொருத்தமான சொற்றொகை கொண்ட கருத்தாடல் களத்திற்கு பொருத்தமான சொற்களின் குழுமங்களும் உள்ளன. ஆனால் அமைப்பு என்ற நிலையில் இவற்றிற்குச் சிறப்பான பண்புகள் இல்லை.

கணினி வழி சொற்களஞ்சிய உருவாக்கம்

சொற்களஞ்சியம் (thesaurus) உருவாக்குவதற்குக் கணிப்பொறியைப் பயன்படுத்தலாம். ஒத்தல் (similarity), உள்ளடங்கல் (Inclusion), மேலுறல்

(Overlapping), வேறுபடல் (Disjunction) என்ற தர்க்கப் பொருண்மை உறவுகளின் அடிப்படையில் சொற்களை அவற்றின் படிநிலை அமைப்பும் பொருண்மையும் வெளிப்படும் விதத்தில் பொருண்மைக் களங்களாகப் பகுத்துத் தருவது சொற்களஞ்சியமாகும். கணிப்பொறி உதவியுடன் இச்சொற்களஞ்சியத்தைக் குறைந்த உழைப்பில் குறுகிய காலக் கட்டத்தில் உருவாக்கிவிடலாம். கணிப்பொறியில் பயன்படுத்தத்தக்க விதத்தில் மின் சொற்களஞ்சியத்தை உருவாக்கலாம். இத்தகைய மின்சொற்களஞ்சியம் மொழிபெயர்த்தல், மொழி கற்பித்தல், சொற்பொருண்மை மயக்க நீக்கம், தகவல் மீட்பு போன்ற செயற்பாடுகளுக்குப் பெரிதும் பயன்படும். தமிழுக்கான மின் சொற்களஞ்சியம் ஒன்றைத் தமிழ்ப் பல்கலைக்கழக மொழியியல் துறையும் கணிப்பொறி அறிவியல் துறையும் இணைந்து உருவாக்கியுள்ளன (இராசேந்திரன் மற்றும் பாஸ்கரன் (2006). இதன் தரவுத்தளம் இராசேந்திரனின் (2001) தற்காலத் தமிழ்சொற்களஞ்சியத்தைப் பயன்படுத்தி உருவாக்கப்பட்டுள்ளது. இச்சொற்களஞ்சியத்தின் தரவுத்தளம் நைடாவின் (1975) பொருண்மைக்கூறாய்வு அடிப்படையில் உருவாக்கப்பட்டது. அவர் பைபிள் மொழிபெயப்புக்காக சொற்களை மெய்படுத்தும் செய்யும் கருத்துருக்களைப் படிநிலை அமைப்பாக பொருட்புல வகைப்பாடு செய்து தந்துள்ளார். (இதில் தரப்பட்டுள்ள சொற்றொகைப் பொருண்மை அமைப்பு காட்சி மூலப்பொருண்மையியல் உருவாக்கத்தின் கீழ் விளக்கமாகத் தரப்பட்டுள்ளது.) கீழ்வருவது இதன் சுருக்கமான கிளைப்படமாகும்.



நைடாவைப் (1975) பின்பற்றி சொற்களால் மெய்ப்படுத்தம் செய்யப்படும் கருத்துருக்கள் (concepts) பருமைகள் (entities), நிகழ்வுகள் (events), அருவங்கள் (abstracts), தொடர்பிகள் (relationals) என நன்காகப் பகுக்கப்பட்டுள்ளன. பருமைகள் அருவமல்லாப் பருமைப்பெயர்களை உள்ளடக்கும். நிகழ்வுகள் வினைகளையும் வினைப்பெயர்களையும் உள்ளடக்கும்; அருவங்கள் அருவப்பெயர்களையும் பெயரடைகளையும் வினையடைகளையும் மெய்ப்படுத்தம் செய்யும்; தொடர்பிகள் பின்னருபுகளையும் பிற செயற்பாட்டுச் சொற்களையும் (functional words) மெய்ப்படுத்தம் செய்யும். இப்பொருண்மைக் களங்களில் உட்படுத்தப்படும் சொற்கள் ஒருபொருப்பொருள் பன்மொழியம்(synonymy), உள்ளடங்கு மொழியம்-உள்ளடக்கு மொழியம் (hyponymy-hypernymy), சினைமொழியம்-முழுமொழியம் (meronymy-holonymy), எதிர்மொழியம் (antonymy) போன்ற பொருண்மை அல்லது சொல்சார் உறவுகளால் தொடர்பு படுத்தப்பட்டுள்ளன. பயனர் இடைமுகத்தின் மூலம் வேண்டிய சொல்லை அதில் உள்ளீடு செய்து அதன் ஒருபொருள் பன்மொழிகளையோ உள்ளடங்கு மொழியையோ உள்ளடக்கு மொழியையோ சினை மொழியையோ முழுமொழியையோ எதிர்மொழியையோ தரவுத்தளத்திலிருது பெறலாம்.

=====
Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

4. தமிழ்க் காட்சி மூலப்பொருண்மையியல் சொற்களஞ்சிய உருவாக்கம்

நைடா (1975) முன்வைத்த பொருளின் பொருண்மைப் பகுப்பாய்வின் கோட்பாடுகள், மரபு அடிப்படையிலான நிகண்டு மற்றும் அரிஸ்டாட்டிலியன் வகை மற்றும் இனங்கள் பற்றிய கொள்கைகளின் அடிப்படையில் 1990ஆம் ஆண்டில் தமிழிற்கான ஒரு சொற்களஞ்சியம் இராசேந்திரனால் உருவாக்கப்பட்டு 2001இல் வெளியிடப்பட்டது (இராஜேந்திரன், 2001). அதைத் தொடர்ந்து, தமிழிற்கான ஒரு மின்னணு சொற்களஞ்சியம் முயற்சிக்கப்பட்டு 2006ஆம் ஆண்டில் தமிழ் மின்னணு சொற்களஞ்சியம் என்ற புத்தகம் வெளியிடப்பட்டது (இராஜேந்திரன் மற்றும் பாஸ்கரன், 2006). தமிழிற்கான சொல்வலை தயாரித்தல் (2001-2003; 2007-2015) திராவிடச் சொல்வலை உருவாக்கத்தின் ஒரு பகுதியாக இராசேந்திரனை முதன்மை ஆய்வாளராகக் கொண்டு அவரது மேற்பார்வையில் தமிழ்ப் பல்கலைக்கழகத்தின் இந்திய அரசின் மனிதவள மேம்பாட்டுத் துறை மற்றும் மின்னணு மற்றும் தகவல் தொழில்நுட்பத் துறையிலிருந்து பெறப்பட்ட நிதி உதவியுடன் கட்டமைக்கப்பட்டு ஒரு நிலையை அடைந்துள்ளது. இதன் பின்னர் இரண்டாயிரது பதினைந்து பதினாறுகளில் சென்னையிலுள்ள தமிழ் இணையக் கல்விக்கழகத்தின் (Tamil Virtual Academy, Chennai) நிதி நல்கையில் இராஜேந்திரன் (இராஜேந்திரன், 2001) உருவாக்கிய சொற்களஞ்சியத்தில் தரப்பட்டுள்ள மூலப்பொருண்மையியல் (ontology) அடிப்படையில் அமிர்தா பல்கலைக்கழகத்தில் கணினிசார் தொழில்நுட்பம் மற்றும் வலைப்பின்னல் மையத்தில் (Centre for Computational Engineering and Networking) நடைபெற்ற “தமிழ் மொழி கற்பித்தல் மற்றும் கற்றலுக்கான கணினி கருவிகள்”

ஆய்வுதிட்டத்தின் ஒரு பகுதியாகத் தமிழ் காட்சி மூலப்பொருண்மையியல் சொற்களஞ்சியம் (Rajendran 2016, 2018) உருவாக்கப்பட்டு அந்நிறுவனத்தின் இணையத்தில் பதிவேற்றம் செய்யப்பட்டுள்ளது. இச்சொற்களஞ்சியம் மேற்சொன்ன மின்சொற்களஞ்சியம், சொல்வலை இவற்றின் அடிப்படையில் அமையும்.

மெய்ப்பொருள் மூலாய்வின் வரையறை விளக்கம்

மூலப்பொருண்மையியல் ஆய்வு ஒரு பகிர்ந்துகொள்ளப்பட்ட கருத்துருவாக்கத்தின் முறையான, வெளிப்படையான சிறப்பீடாகும். கருத்துவாக்கம் என்பது ஒரு நடத்தையின் தேவையான கருத்துருக்களின் கண்டுபிடிப்பால் உருவாக்கப்பட்ட ஒரு அருவ மாதிரியாகும். கருத்துருக்களும் அவற்றிற்கிடையேயுள்ள உறவுகளும் அவற்றின் பயன்பாட்டின் கட்டுப்பாடுகளும் வெளிப்படையாக வரையறை விளக்கம் செய்யப்பட்டுள்ளது. “முறையானது” (formal) என்பது மூலப்பொருண்மையியல் ஆய்வு இயந்திரத்தால் படிக்க இயலுவது மற்றும் இயற்கை மொழியின் பயன்பாட்டை விலக்கி வைக்கின்றது. எடுத்துகாட்டாக, மருத்துவ பொருட்புலத்தில், கருத்துருக்கள் யாவும் நோய்களும் அறிகுறிகளும் ஆகும்; அவைகளுக்கிடையேயுள்ள உறவு இயல்பானது மற்றும் நோய் தானே காரணமாகாது என்பது கட்டுப்பாடாகும். மூலப்பொருண்மையியல் ஆய்வு ஒரு “பகிர்ந்துகொள்ளப்பட்ட கருத்துருவாக்கம்” என்பது ஒரு குழுவின் பயன்பாட்டிற்குரிய ஒருமித்த அறிவை உருப்படுத்தம் செய்வதை நோக்கமாகக் கொண்டது என்று கூறும். இலட்சிய அளவில் மூலப்பொருண்மையியல் ஆய்வு அதன் பயன்பாட்டிலிருந்து சுதந்திரமாகவும் உலகளவில் பகிர்ந்துகொள்ள இயலும் வழியிலும் அறிவைக் கைகொள்ளும்; ஆனால் நடைமுறையில் வேறுபட்ட பணிகளும் பயன்பாடுகளும் மெய்ப்பொருள் மூலாய்வில் அறிவின் வேறுபட்ட உருப்படுத்தத்தை வேண்டும்.

மூலப்பொருண்மையியல் ஆய்வு சில வேளைகளில் வகைப்பாட்டியலுடன் குழப்பிக்கொள்ளப்படும். வகைப்பாட்டியல் என்பது ஒரு பொருட்புலத்திலுள்ள தரவின் வகைப்படுத்தலாகும். இரு முக்கியமான சூழல்களில் அவைகள் வேறுபடும்:

1. ஒரு மூலப்பொருண்மையியல் ஆய்வு கருத்துருக்களுக்கிடையே உறவுகளையும் கட்டுப்பாடுகளையும் உட்படுத்தியுள்ளதால் வளமான அகவயமான அமைப்பைக் கொண்டிருக்கும்.
2. மூலப்பொருண்மையியல் ஆய்வு பொருட்புலத்தின் அறிவைப் பற்றிய சில ஒருமித்த கருத்தை உருப்படுத்தம் செய்வதாக வலியுறுத்தும். இந்த ஒருமித்த கருத்து அறிவின் திட்டமிடப்பட்ட (இலக்கு) பயன்பாட்டாளர்களுக்கிடையில் ஆகும். எ.கா. சில நோய்களுக்கு வேண்டி ஒரு மருத்துவ மனையின் மூலப்பொருண்மையியல் ஆய்வை மருத்துவர்கள் பயன்படுத்துவது; வரலாற்றுக் கலையுடன் தொடர்புடைய மூலப்பொருண்மையியல் ஆய்வை கலைஞர்கள் பயன்படுத்துவது.

மூலப்பொருண்மையியல் ஆய்வுகள் அவை உருப்படுத்தம் செய்யும் அறிவைப் பற்றிய பொதுவான உடன்பாட்டின் ஒரு வடிவத்தை உருப்படுத்தம் செய்யவதை நோக்கமாகக் கொண்டுள்ளதால், அவைகள் பெரும்பாலும் வேறுபட்ட மக்களை உட்படுத்தி கூட்டுறவு செயல்முறையில் உருவாக்கப்படுகின்றன. மூலப்பொருண்மையியல் ஆய்வுகள் அவை கொண்டிருக்கும் கொள்கைகளின் பொதுமைத்தன்மையின் அளவு (degree) அடிப்படையில் பிரிக்கப்பட்டுள்ளது.

நைடாவின் சொற்றொகை வகைபாட்டு அடிப்படையில் மூலப்பொருண்மையியல் உருவாக்கம்

நைடாவின் பொருண்மையமைப்பு/ பொருட்புலப் பாகுபாடு அடிப்படையில் தமிழ்ச் (Nidat 1975a) சொற்றொகை மூலப்பொருண்மையியல் கோட்பாடு

அடிப்படையில் காட்சி மூலப்பொருண்மையியல் சொற்களஞ்சியத்திற்காக அமைப்பாக்கம் செய்யப்பட்டுள்ளது.

நைடா (1975a) குறிப்புப்பொருளின் (Referential meaning) பொருளணுப் பகுப்பாய்வினை ஒரு மொழியின் சொற்றொகையின் பொருண்மை அமைப்பைக் காண்பதற்குப் பயன்படுத்துகிறார். சொல்லுக்கும் அதன் குறிப்புக்கும் (referent) உள்ள தொடர்புதான் குறிப்புப் பொருளாகும். குறிப்பு பொருள் இரண்டு காரணிகளை அடிப்படையாகக் கொண்டது. அவை (உணர்ச்சி பொருளுக்கு எதிரான) புலனறிவுப் பொருளும் (இலக்கணப் பொருளுக்கு எதிரான) குறிப்பு பொருளும் ஆகும். குறிப்புப் பொருள் ஒரு சொல்லின் கட்டமைந்த ஒரு கூட்டம் புலன்கூறுகள் (cognitive features) அடங்கியதாகும். இப்புலன் கூறுகள் ஒரு சொல் மற்றொரு சொல்லிலிருந்து வேறுபடப் பயன்படுகின்றன. ஒரு பரப்பில் (semantic domain) அடங்கும் சொற்கள் தங்களுக்குள் வேறுபட உதவும் இம்முக்கியமான பொருண்மைக் கூறுகளை நாம் கண்டுபிடிக்க வேண்டும். ஒத்த பொருண்மைக் கூறுகளைக் கொண்டிருக்கும் குறிப்புப் பொருளை ஒன்றாக ஒரு பொருள் பரப்பில் அடக்கலாம். இப்பொருள் பரப்பின் தேவை (Relevancy), பருப்பம் (largeness), படிநிலை அமைப்பில் (காண்க பக்கம்:) அதன் செயற்பாட்டு நிலை மட்டம் (level of function) ஆகியன மொழியின் மொத்த பொருண்மைக் கட்டமைப்பைப் பொறுத்திருக்கின்றன.

ஒன்றுக்கொன்று தொடர்புடைய பொருள் தொகைகளின் பொருண்மைத் தொடர்புகளின் விளக்கம் இரண்டு வேறுபட்ட வகையான உறவுகளை எடுத்துக் கொள்ள வேண்டும். அவையாவன வெவ்வேறு சொற்களுக்கு இடையே உள்ள தொடர்புடைய பொருள்கள் மற்றும் ஒரு சொல்லின் தொடர்புடைய பொருள்களுக்கு இடையே உள்ள உறவு. ஒரு பொருள் பரப்பில் அடங்கும் வேறுபட்ட சொற்களின்

பொருள்களை நான்கு வழிகளில் உறவுபடுத்தலாம். அவை உள்ளடங்கல் (inclusion), மேலுறல் (overlapping), துணைநிலைப்படுதல் (complementation), அண்மைப்படல் (contiguity), ஆகியன. உள்ளடங்கல் படிநிலைத் தொடரில் தங்களுக்குள் தொடர்புள்ள மற்றும் அவை அடங்கியுள்ள பொருளிலிருந்து குறைந்தது ஒரு வேறுபடுத்தும் பொருளணு கூடுதலாக உள்ள பொருள்கள் அடங்கிய தொகுதியை உருவாக்கும் (எ.கா. விலங்கு: பாலூட்டி). மேலுறல் ஒரு படிநிலை மட்டத்தில், மரபுப்படி ஒரு பொருள் பன்மொழிகள் என்று அறியப்படும் மற்றும் அவற்றின் மேலுறல் தொகுதியை உருவாக்கும் (எ.கா கொடு: அளி, காய்ச்சல்: சுகக்கேடு). துணைநிலை பொருள்களின் தொகுதிகள் மூன்று முக்கிய வழிகளில் தொடர்புடையன. அவை எதிர்நிலைச் சொற்கள் (opposites) (எ.கா. பெரிது: சிறிது) திருப்புப் பொருள் நிலைச் சொற்கள் (reversies) (எ.கா. கட்டு: அவிழ்) மறுதலைச் சொற்கள் (conversies) (எ.கா. வாங்குதல்: விற்றல்) ஆகியன. ஒரு படிமட்டத்தில் கூடுதல் பொருள் ஒப்புமை உள்ள அதே சமயம் குறைந்தது ஒரு முக்கிய பொருளணுவால் குறிப்பிடத்தக்க விதத்தில் வேறுபடும் பொருள்கள் ஒரு அண்மைப்பட்ட தொகுதியை உருவாக்கும் (எ.கா நாகாலி: பெஞ்சு: ஸ்டூல்: சோபா).

ஒரு சொல்லின் வேறுபட்ட பொருள்களை நான்கு வழிகளில் தொடர்புபடுத்தலாம்: ஆக்கம் (derivation) மாற்றீடு (replacement) உருவகநீட்சி (figurative extension) மற்றும் புற எல்லைக் கோவை (peripheral clustering) அடிப்படைப் பொருளணுக்களை அதனின்றும் பெறப்பட்ட வேறு பொருள் பரப்பைச் சார்ந்த பொருளில் அடங்குவது ஆக்கமாகும். (எ.கா காதல்: காதலன்). அடிப்படைப் பொருளின் குறைந்தது ஒரு பொருளணுவால் பொருள் பரப்பில் பெரிய மாற்றம் விளைவிக்காமல் இடமாற்றம் செய்வது மாற்றீடாகும் (எ.கா. ராஜா: ராணி). பெரிதும் வேறுபட்ட பொருள் பரப்புகளுல் அடங்கும். அடிப்படைப் பொருளுக்கும் உருவகப்பொருளுக்கும் உள்ள தொடர்பைச் சில துணைப் பொருளணுக்கள்

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

(supplementary components) அல்லது திரும்பி பொருள் கொள்ளப்பட்ட பொருளணுக்கள் (reinterpreted diagnostic components) வாயிலாக நிறுவது பொருளின் உருவக நீட்சியாகும் (எ.கா. நரி: தந்திரசாலி) மத்தியப் பொருளுடன் (central meaning) தொடர்புடைய அல்லது அத்துடன் சிலப் பொதுப் பொருளணுக்களைக் கொண்டு தங்களுக்குள் தொடர்புடைய புற எல்லைப் பொருள்களைக் கோவைப்படுத்துதல் புற எல்லைக் கோவையாகும். (எ.கா கெட்டியான தோல், கெட்டியான பால் இவற்றில் வரும் கெட்டி).

நைடா (1975b:152) பின்வருமாறு கூறுகிறார்: பொருள் உணர்த்தும் பலவகை சொற்சேர்க்கையின் ஆய்வு நம்மை நான்கு அடிப்படை பொருண்மைக் குறியீடுகளின் வகுப்புகளை இனங்கண்டுகொள்ளத் தூண்டும். இவ்வகுப்புகளில் அடங்குபவை: 1. குதிரை, மரம், இலை, சூரியன், மலை, போன்ற 'பருப்பொருள்கள்' (objects), 2. ஆடுதல், நடத்தல், வெட்டுதல், தள்ளுதல், போதல், இருத்தல், புரள்தல், போன்ற 'நிகழ்வுகள்' (events), 3. அ) சிவப்பு, நீலம், சின்னது, பல, ஒன்று, இரண்டு, (பருப்பொருள்களின் அருவங்கள்), ஆ) விரைவு, மெல்ல, ஒருதடவை, இரண்டு தடவை, அடிக்கடி, (நிகழ்வுகளின் அருவங்கள்), இ) கடினம் (எ.கா கடினமான பாறை, கடினமான வேலை) மிருது (எ.கா மிருதுவான கோட்டு, மிருதுவான வேலை) (பருப்பொருள்கள் அல்லது நிகழ்வுகள் என்பதன் அருவங்கள்) அல்லது ஈ) மிக (எ.கா மிகப்பல) ரொம்ப (எ.கா ரொம்ப சின்ன) வெகு (எ.கா வெகு விரைவு) (அருவங்களின் அருவங்கள்) இவற்றைத் தொடர்புபடுத்த உதவும் 'தொடர்பான்கள்' (rationals), எ.கா அ) பருப்பொருளுடன் பருப்பொருள்: ஜானும் பில்லும் , வீட்டுள்ள மனிதர்கள், மேஜை மேலுள்ள புத்தகம்: ஆ) பருப்பொருளுடன் நிகழ்வுகள்: பறக்கும் மனிதர்கள், வேட்டை நாய்களின் குரைப்பு ; இ) நிகழ்வுடன் நிகழ்வு: ஓட்டமும் குதிரையும், நீச்சலில் வெற்றி; ஈ) பருப்பொருளுடன்

அருவம்; பொருளின் அழகு, ரோஜாவின் நிறம்; உ) அருவத்துடன் அருவம்: மோசமும் சோம்பலும், பரிசுத்த அழகு. இந்த நான்கு அடிப்படைப் பொருண்மை வகுப்புகளும் எல்லா மொழிகளிலும் காணப்படுவதுடன் எல்லா வகையான சொல்லுருக்களுக்கு இடையே உள்ள பொருள் தொடர்பை ஆய்வதற்கும் அடிப்படையாக அமையும் என்றும் தோன்றுகின்றது.

நைடா கிரேக்கப் புதிய வேதாகமத்திற்கு ஒரு புதிய அகராதி உருவாக்க முனைந்து அதை நிறைவேற்றியுள்ளார். இது பொருள்களைச் சொல் அலகுகளின் அகர வரிசை நிரலில் அல்லாமல் பொருள் பரப்பின் அடிப்படையில் முறைப்படுத்தும். வெவ்வேறு பொருண்மை அலகுகளின் தொடர்பான அர்த்தங்களின் அடிப்படையில் பொருள் பரப்புகள் பகுக்கப்பட்டுள்ளன. தேவையான பொருணுக்களைக் காணுவதில் ஒரு பொருள் பரப்புகள் பகுக்கப்பட்டுள்ளன. தேவையான பொருள்களைக் காணுவதில் ஒரு பொருள் பரப்பின் குறிப்பு நுட்பம் பொதுவாக எந்த ஒரு பொருள் தொகுதியும் தம்முள் அடங்கியுள்ள பொதுவான பொருளணுக்களின் எண்ணிக்கையோடு தகவுப்பொருத்தம் கொண்டிருக்கும். இந்த அகராதி முக்கியமாக மொழிகளுக்கு இடையே உள்ள பொருளணு அடிப்படையிலான வேறுபாடுகளை அறிந்த பைபிள் மொழிபெயர்ப்பாளர்களுக்காகத் திட்டமிடப்பட்டதாகும். நாம் முன்னர் கண்ட சொல் உறவுகளின் அடிப்படையில் பொருளணு ஆய்வைப் பயன்படுத்தி அவர் கிரேக்க மொழியின் சொற் கட்டமைப்பைத் தர முயன்றுள்ளார். கீழே தரப்பட்டுள்ளது நைடா (1975a:176-186) தந்துள்ள வகைப்பாட்டின் சுருக்கமாகும்.

1. பருப்பொருள்

எ. விலங்குகளல்லாதவை

1. இயற்கையானவை

எ. பூகோளம் சம்பந்தப்பட்டவை

பி. இயற்கை சாதனங்கள்

சி. தாவரங்கள் மற்றும் தாவரப்பொருட்கள்

2. உருவாக்கிய அல்லது கட்டப்பட்ட பொருள்கள்

எ. கலைப்பொருள்

பி. பதம் செய்யப்பட்ட பொருள்கள்: உணவுகள் மற்றும் வாசனை பொருள்கள்

சி. கட்டிடங்கள்

பி. விலங்குப்பொருள்கள்

1. விலங்குகள், பறவைகள், பூச்சிகள்

2. மனிதர்கள்

3. இயற்கைக்கு அப்பாற்பட்ட சக்திகள் அல்லது உயிர்கள்

நிகழ்வுகள்

பௌதீகம் தொடர்பானவை, உடற்கூறு தொடர்பானவை, உணர்வு தொடர்பானவை, உணர்ச்சி தொடர்பானவை, அறிவு தொடர்பானவை, கருத்துப்பரிமாற்றம், சேர்க்கை, கட்டுப்பாடு, சலனம், தாக்கம், கைமாற்றம், தொடர்ச்சியான சலனங்கள் அல்லது செயல்கள் உள்ளடங்கிய பல்கூட்டச் செயல்கள்.

அருவங்கள்

காலம், தூரம், பருமை, விரைவு, சூடு, நிறம், எண், நிலை, மதப்பண்பு, கவர்ச்சி, வயது, உண்மை-பொய், நல்லது-கெட்டது, பலம், ஆரோக்கியம் நிலை.

தொடர்பான்கள்

இடம் தொடர்பானவை, காலம் தொடர்பானவை, கட்டல், தர்க்கம், தொடர்பானவை.

மேற்சொன்ன வகைப்பாடு குறிப்புப் பொருள்களை அடிப்படையாகக் கொண்டது. சொற்பரப்பு வகுப்புகளுக்கும் இலக்கண வகுப்புகளுக்கும் ஒன்றுக்கொன்று

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

நிகரான பொருத்தம் இல்லை. இருப்பினும் அவை தம்மில் சில இசைவுகள் இருக்கத்தான் செய்கின்றன. ஏனென்றால் ஏதோ ஆழ்நிலை அமைப்பு மட்டத்தில் பருப்பொருட்கள் பெயர்களாலும் நிகழ்வுகள் வினைகளாலும் அருவங்கள் அடைகளாலும் தொடர்பான்கள் பல வேறுபட்ட கூறுகளாலும் (இடைச் சொற்கள், வேற்றுமை ஒட்டுகள், சொல் நிரல் போன்றவை) குறிக்கப்படுகின்றன. மேற்கண்ட வகைப்பாடு நிச்சயமாக ஒரு தெளிவான தர்க்க நோக்கால் அமைந்ததாகும். இருப்பினும் இதே சிந்தையை அடிப்படையாகக் கொண்டு சிறிது வேறுபட்ட ஒரு வகைப்பாட்டைச் செய்ய இயலும்.

நைடா உலகப்பொருண்மை வகைப்பாட்டை அதிகமான மாற்றங்கள் செய்யாமல் தமிழுக்கு ஏற்றுக் கொள்ளலாம் என்றாலும் இவ்வாறு செய்வதிலும் பல சிக்கல்கள் வரத்தான் செய்கின்றன. நைடாவே தன் வகைப்பாட்டின் குறைப்பாடுகள் பற்றிக் குறிப்பிடுகிறார் (Nida, 1975:187): “இந்தச் சொற்பரப்புகளின் பட்டியல் தற்காலிகமானதாகும். அகராதிச் சொற்களஞ்சியத்தை (Dictionary-thesaurus) உருவாக்கும் குழுக்கள் பதிப்புவேலையைத் தொடர்ந்து கொண்டிருப்பதால் இந்த வகைப்பாடு மாற்றத்திற்கு உரியதாகும். மேலும், இதில் கொடுக்கப்பட்டுள்ள பட்டியல் முழுமையானதல்ல. ஏனென்றால் இது ஒரு சில சொற்பரப்புகளைத்தான் பட்டியல் படுத்தியுள்ளது. பெரும்பாலும் ஒரு சில மாதிரிக்கான பொருள்கள் தாம் கொடுக்கப்பட்டுள்ளன.” (நைடாவின் கிரேக்க அகராதிச் சொற்களஞ்சியம் வெளிவந்துவிட்டது.) நைடாவின் பொருண்மைச் சட்டகத்தில் பொருந்தும் படி தமிழ்ச் சொற்களை வகைப்பாடு செய்வது சில இடங்களில் கடினமாக அமைகிறது. ஒரு குறிப்பிட்ட சொல் எந்தப் பொருள் பரப்பைச் சார்ந்தது என்பதை முடிவு செய்வதில் அடிக்கடி தெளிவின்மை ஏற்படுகிறது. நைடாவே பின்வருவனவற்றை முக்கியமான

சிக்கல்களாக காட்டுகிறார். 1. சொல்லாக்கப் பொருளுக்கும் அதன் அடிப்படைக்கும் உள்ள உறவு, 2. சொந்திரிபு வேற்றுமைகளைக் கையாளுதல், 3. ஒரு சொல்லின் வேறுபட்ட பொருள்கள் வேறுபட்ட மட்டங்களில் வருதல், 4. சொற்பரப்புகளுக்குப் பொதுப்பொருள் வருதல் அல்லது வராமை, 5. படிநிலை வகைப்பாட்டின் கட்டமைப்பு, 6. சொற்பரப்புகளுக்கு இடையே உள்ள எல்லைகள் இச்சிக்கல்களை மனதில் கொண்டு சொற்களஞ்சியப் பொருட்புல வகைப்பாடு செய்யப்படவேண்டும்.

நடாவின் சொற்றொகைப் பாகுபாடு அடிப்படையில் தமிழ்ச்சொற்றொகை பகுக்கப்பட்டு சொல்வலையில் விளக்கப்பட்டுள்ள சொல்சார் உறவுகள் மற்றும் பொருண்மை உறவுகள் அடிப்படையில் உறவுபடுத்தப்பட்டுள்ளன. இதுதான் இச்சொற்களஞ்சியத்தின் தரவுதளமாகும்.

காட்சி மூலப்பொருண்மையியல் சொற்களஞ்சிய உருவாக்கம்

ஒரு பயனரின் வினவலுக்காக கொடுக்கப்பட்ட மிகப்பெரிய தரவுத்தளத்திலிருந்து மூலப்பொருண்மையியல் கிளையமைப்பை உருவாக்குவது என்பது ஒரு புத்திசாலித்தனமான தகவல் மீட்டெடுப்பு மற்றும் காட்சிப்படுத்தல் ஒழுங்கமைப்பாகும். இந்த வகைக் கருவி தரவு உருப்படுத்தம் மற்றும் அறிவுப் பிரித்தெடுத்தல் (பெறுதல்) செயல்முறைக்கு ஒரு சக்திவாய்ந்த வழியை வழங்குகிறது. பயனர் கணினியில் தேடவேண்டிய வார்த்தையை உள்ளீடு செய்து தரவுத்தளத்திலிருந்து தகவல்களைப் பெறலாம். தகவல் ஒரு காட்சி உருப்பத்தமாகக் கொடுக்கப்பட்டுள்ளது. இந்த நோக்கத்திற்காக நன்கு கட்டமைக்கப்பட்ட தரவு தளம் தேவை. பின்வரும் அட்டவணை சொற்கள் அல்லது சொற்பொருள் இவைகளுக்கு இடையில் நிறுவப்பட்ட பல்வேறு வகையான அர்த்த உறவுகளையும்

மூலப்பொருண்மையியல் கிளை அமைப்பு உருவாக்குவதற்குத் தரவுத்தளத்தில் அவற்றைக் குறிக்கும் தனித்துவமான வழியை விளக்குகிறது. தற்போதைய காட்சிச் சொற்களஞ்சியத்தில் பின்வரும் எழு உறவுகளே நிறுவப்பட்டுள்ளன. இவ் உறவுகளின் எண்ணிக்கை எதிர்காலத்தில் அதிகரிக்கப்படும்.

பொருண்மை மற்றும் சொல்சார் உறவுகள்	விளக்கமும் எடுத்துக்காட்டும்	உறவை நிறுவும் குறிப்பான்கள்
இணைமொழியம் (ஒருபொருள்பன்மொழியம்)	ஒரே பொருண்மையைக் குறிக்கும் சொற்களுக்கு இடையே உள்ள உறவு. எ.கா. புத்தகம், நூல் என்பவைகளுக்கு இடையே உள்ள உறவு.	< >
வகைமொழியம்	உள்ளடங்கு மொழிக்கும் உள்ளடக்கு மொழிக்கும் இடையே உள்ள உறவு. நாய், பூனை போன்ற உள்ளடங்கு மொழிகளுக்கும் விலங்கு என்ற உள்ளடக்கு மொழிக்கும் இடையில் உள்ள உறவு.	{ }
பகுதிமொழியம்	சினைமொழிக்கும் முழுமொழிக்கும் உள்ள உறவு. கை,கால், தலை என்ற	/

	சினைமொழிகளுக்கும் உடல் என்ற சினை மொழிக்கும் உள்ள உறவு.	
எதிர்மொழியம்	எதிர்மொழியம்	% *
பால்மொழியம்	பலர்பால் ஒருமை வடிவிற்கும் ஆண்பால் ஒன்றன்பால் மற்றும் பொண்பால் ஒன்றன்பால் வடிவிற்கும் உள்ள தொடர்பு. எடுத்துக்காட்டாக வேலைக்காரர் என்பதற்கும் வேலைக்காரன், வேலைக்காரி என்பதற்கும் உள்ள உறவு	\$ #
வினைப்பெயராக்கமொழியம்	வினைக்கும் அதிலிருந்து உருவாக்கப்பட்ட பெயர்ச்சொற்களுக்கும் இடையிலுள்ள ஆக்க உறவு வினைப்பெயராக்கமொழிய உறவு என்று அழைக்கப்படும். எடுத்துக்காட்டாக படி என்ற வினைக்கும் படிப்பு, படித்தம், படித்தல் என்ற	~ `

	வினைப்பெயர்களுக்கும் உள்ள உறவு.	
அடைமொழியம்	பெயருக்கும் அதிலிருந்து ஆக்கப்பட்ட பெயரடைகளுக்கும் வினையடைகளுக்கும் இடையே உள்ள உறவு. எத்துக்காட்டாக அழகு என்ற பெயருக்கும் அதிலிருந்து உருவாக்கப்பட்ட அழகான என்ற பெயரடை மற்றும் அழகாக என்ற வினையடை இடையே உள்ள உறவு.	^!

ஒருபொருள் பன்மொழியம்

ஒரே பொருளைக் கொண்ட அல்லது பொருளின் ஒரே பொருண்மைக்கூறுகளைப் பகிர்ந்து கொள்ளும் சொற்களுக்கு இடையே உள்ள உறவு ஒருபொருள் பன்மொழியம் எனப்படும்; இவ் உறவைக் காட்டும் சொற்கள் ஒழுபொருள் பன்மொழிகள் எனப்படும். ஒத்த பொருள் என்பது சம்பந்தப்பட்ட சொற்கள் அர்த்த வெளிப்படுத்தத்தில் ஒரே மாதிரியாக இருக்க வேண்டும் என்று அர்த்தமல்ல; கருத்தப்பட்ட சூழல்களில் வாக்கியத்தின் அர்த்தத்தில் எந்த வித்தியாசமும் இல்லாமல் ஒன்றுக்கொன்று இடம்பெயர்க்க இயலும் சொற்கள் ஒருபொருள் பன்மொழிகளாகும்.

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

உதாரணமாக, சொற்களை எடுத்துக் கொள்ளுங்கள். எடுத்துக்காட்டாக *நான் புத்தகம் படித்தேன்* என்பதும் *நான் நூல் படித்தேன்* என்பதும் ஒரே பொருண்மையைத் தருவதால் *புத்தகம்* என்பதும் *நூல்* என்பதும் ஒருபொருள் பன்மொழிகளாகும். ஒருபொருள் பன்மொழியச்சொற்கள் தரவுத்தளத்தில் '< >' என்ற அடைப்புக் குறிக்குள் தரப்பட்டிருக்கும். பின்வருவது ஒரு எடுத்துக்காட்டாகும்.

சூரியன் < ஞாயிறு, ஆதவன், பகலவன், திவாகரன், ஆதித்தன், கதிரவன், கதிரோன், கிரணன், செங்கதிரோன், செங்கதிர், வெங்கதிரோன், வெங்கதிர், வெஞ்சுடர் ஆயிரங்கதிரோன், ஆயிரங்கிரணன், உலகநேந்திரன், உலகப்பாந்தவன் >; { எழுஞாயிறு, ஏறுஞாயிறு; சாய்ஞாயிறு, இறங்குஞாயிறு; இளஞ்சூரியன், < பாலசூரியன் >; }.

உள்ளடங்குமொழியம்-உள்ளடக்குமொழியம் (Hyponymy and Hypernymy)

உள்ளடங்கு மொழி என்பது ஒருசொல் மற்றொருசொல்லின் கீழ் பொருண்மை அடிப்படையில் உட்படுத்தப்படுவதாகும். உட்படும் சொல் உள்ளடங்கு மொழியம் என்றும் உட்படுத்தும் சொல் உள்ளடக்கு மொழியம் என்றும் அழைக்கப்படும். எடுத்துக்காட்டாக நாய், பூனை என்ற உள்ளடங்குமொழிகள் விலங்கு என்ற உள்ளடக்கு மொழியின் கீழ் பொருண்மை அடிப்படையில் உட்படுத்தப்படும். நாய், பூனை என்பன விலங்கு என்பதன் இணை உள்ளடங்கு மொழிகளாகும். தரவுத்தளத்தில் இணை உள்ளடங்கு மொழிகள் உள்ளடக்கு மொழிகளுக்கீழ் மீசை அடைப்புக்குறிகளுக்குள் தரப்பட்டுள்ளன. "{ }". பின்வருவது எடுத்துக்காட்டாகும்.

சப்தமண்டலம்: {வாயுமண்டலம்; வருணமண்டலம்; சந்திரமண்டலம்;
சூரியமண்டலம்; நட்சத்திரமண்டலம்; அக்கினிமண்டலம்; திரிசங்குமண்டலம்; }

நவகிரகம்: {சூரியன்; சந்திரன்; செவ்வாய்; புதன்; வியாழன்; சூக்கிரன்; சனி, ராகு, கேது}

சினைமொழியம்-முழுமொழியம் (Meronymy-Holonymy)

சில சொற்கள் அவை மற்றொரு சொல்லின் உறுப்பாக வரும் சொல் என அகராதியில் விளக்கப்பட்டுள்ளதைக் காணலாம். உறுப்பாய் வரும் சொற்கள் சினைமொழி என்றும் முழுமையாய் (உடலாய்) வரும் சொல் முழுமொழி என்றும் அழைக்கப்படும். எடுத்துக்காட்டாக *கை*, *கால்* என்ற சினைமொழிகள் *உடல்* என்ற முழுமொழியுடன் உறவு படுத்தப்பட்டிருக்கும். தரவுதளத்தில் சினைமொழிகள் அவற்றின் முழுமொழிக்குக் கீழ் “/ |” என்ற குறியீடுகளுக்குள் தரப்பட்டிருக்கும். பின்வருவது எடுத்துக்காட்டாகும்.

உடல்: /தலை; கழுத்து; கை,கால்; மார்பு; வயிறு; புருடம்|

காற்றுமண்டலம்: /வெளிமண்டலம்; வெப்பமண்டலம், உஷ்ணமண்டலம்; படலம்: அயனிப்படலம்; ஓசோன்படலம்; |

எதிர்மொழியம்

இரண்டுசொற்கள் எதிரிடையான பொருள்கள் தருவதாகக் கருதப்படுமானால் அவைகளுக்கு இடையே உள்ள உறவு எதிர்மொழிய உறவு எனப்படும். அவைகள் ஒன்றுக்கொன்று எதிர்மொழிகளாக அமையும். *பெரிய* என்பது *சிறிய* என்பதன் எதிர்மொழியாகும் அல்லது *பெரிய*, *சிறிய* என்பன எதிர்மொழிகளாகும். எதிர்மொழிகள் தரவுதளத்தில் “% *” என்ற குறியீடுகளால் இணைக்கப்பட்டிருக்கும். பின்வருவது எடுத்துக்காட்டாகும்.

இருட்டான %வெளிச்சமான*; குட்டியான %நீளமான*; சூடான %குளிர்ச்சியான*

பால்மொழியம்

சில ஆண்பாலுக்கும் பெண்பாலுக்கும் அடையாளப் படுத்தப்பட்ட சில சொற்களுக்கும் அவற்றின் பலர்பாலைக் குறிக்கும் சொற்களுக்கும் இடையே உள்ள உறவு பால்மொழியம் என இங்கு குறிப்பிடப் படுகின்றது. எடுத்துக்காட்டாக வேலைக்காரர் என்பதற்கும் வேலைக்காரன், வேலைக்காரி என்பதற்கும் உள்ள உறவு பால்மொழியம் எனப்படும். தரவுதளத்தில் பலர்பால் சொல்லுக்கும் பால் அடையாளப்படுத்தப்பட்ட சொற்களுக்கும் இடையே உள்ள உறவு பால் உறவு எனப்படும். இவைகளுக்கு இடையே உள்ள உறவு தரவுத்தளத்தில் பின்வரும் குறியீட்டால் உறவுபடுத்தப்படும்: "\$ #". பின்வருவது எடுத்துக்காடாகும்.

இளைஞர்கள்: {இளைஞர் \$ இளைஞன் #, வாலிபர் \$ வாலிபன் #, வாலிபப்பையன், இளவல், இளந்தாரி, இளைஞோர் \$ இளைஞோன் #, காளை, விடலை, இளவட்டம், வயசுப்பையன், வயசுப்பிள்ளை, வயசுப்பிள்ளையாண்டன், இளம்வயதினர், \$இளம்வயதினன்#, இளம்பருவத்தார் \$ இளம்பருவத்தான்#, இளம்பிராயத்தார் \$ இளம்பிராயத்தான்#, பதின்பருவத்தினர் \$ பதின்பருவத்தினன் #, பதினவயதினர் \$ பதினவயதினன் #, வயதுவந்தோர், <வயதுவந்தவர்> \$ வயதுவந்தவன் #, இளந்தை, இளந்தாரி, வாலியன், பாலியன்; இளரத்தம்> }

வினைப்பெயராக்கமொழியம்

வினைக்கும் அதிலிருந்து உருவாக்கப்பட்ட பெயர்ச்சொற்களுக்கும் இடையிலுள்ள ஆக்க உறவு வினைப்பெயராக்கமொழிய உறவு என்று அழைக்கப்படும். எடுத்துக்காட்டாக படி என்ற வினைக்கும் படிப்பு, படித்தம், படித்தல் என்ற

வினைப்பெயர்களுக்கும் உள்ள உறவு வினைப்பெயராக்கமொழியம் எனப்படும். வினைக்கும் அதனுடன் தொடர்புள்ள வினைப் பெயர்களும் தரவுத்தளத்தில் பின்வரும் குறியீடுகளால் உறவுப்படுத்தப்படும்: "~`".

நிகழ்வுகள் குறித்தவை: { நிகழ்; ~ நிகழ்தல்; நிகழ்ச்சி; நிகழ்வு; ` < சம்பவி; ~ சம்பவித்தல்; சம்பவம்; ` விளை; > ~ விளைதல்; விளைவு; ` >நேர்; ~ நேர்தல்; ` பவி; ~ பவித்தல்; ` ஈடேறு; ~ ஈடேறுதல்; ஈடேற்றம்; ` நிலவு; ~ நிலவுதல்; ` }

அடைமொழியம்

பெயருக்கும் அதிலிருந்து ஆக்கப்பட்ட பெயரடைகளுக்கும் வினையடைகளுக்கும் இடையே உள்ள உறவு அடைமொழியம் எனப்படும். எத்துக்காட்டாக அழகு என்ற பெயருக்கும் அதிலிருந்து உருவாக்கப்பட்ட அழகான என்ற பெயரடை மற்றும் அழகாக என்ற வினையடை இடையே உள்ள உறவு அடைமொழியம் ஆகும். பெயர்கள் அதிலிருந்து உருவாக்கப்பட்ட அடைமொழிகளுடன் பின்வரும் குறியீடுகளால் உறவுப்படுத்தப்படும்: "^!". பின்வருவனது எடுத்துக்காட்டாகும்:

நீண்டகாலம்தொடர்பானவை: { நீண்டகாலம்; ^ நீண்டகாலமாக !: நீடுழி, < நெடுங்காலம் ^ நெடுங்காலமாக; ! ரெம்பகாலம்; ^ ரெம்பகாலமாக; ! நீண்டநேரம், < நெடுநேரம், ரெம்பநேரம், நிறையநேரம்; > நீண்டநாள், < நெடுநாள்; ரெம்பநாள், நிறையநாள் >; பல்லாண்டு; நூறாண்டு, < நூற்றாண்டு > }

பயனர் இடைமுகம்

பின்வரும் இடைமுகத்தில் பயனர்கள் வேண்டிய சொல்லை உள்ளீடு செய்து மூலப்பொருண்மையியல் சார் கிளைக் காட்சியைப் பெறலாம். பின்வரும் இடைமுக

5. சொல்வலையும் சொல்வலை உருவாக்கமும்

சொல்வலை

இது ஒரு இணையதளத் தரவு மையமாகும் (Online Data Base). இது சொற்களஞ்சியம் மற்றும் அகராதி என்ற இரண்டின் பண்புகளையும் கொண்டு மிகுந்த திறன் உடையதாய் அமைந்துள்ளது. சொற்கள் தமக்கிடையே உள்ள பொருண்மை உறவுகளால் வலை போல் பின்னப்பட்டிருப்பதை வெளிப்படுத்துவதாய் இது அமையும். சொல்வலையின் மிகப் பரந்த நோக்கம் சொல் தகவல்களைச் சொல்வடிவுகள் அடிப்படையில் அல்லாமல் சொற்பொருண்மை அடிப்படையில் ஒழுங்குபடுத்துவதாகும். இந்த அடிப்படையில் சொல்வலை அகராதியிலிருந்து வேறுபட்டுச் சொற்களஞ்சியத்தையும் உள்ளடக்கும். ரொஜஸ்ட் தெசாரஸ் போன்ற சொற்களஞ்சியங்களும் தமிழுக்கு இராசேந்திரனால் (2001) உருவாக்கப்பட்ட தமிழ்ச் சொற்களஞ்சியமும் சொல்வலை பின்னுவதற்கு முன்னோடியாய் அமையும். இவையெல்லாம் புத்தக வடிவில் அமைந்தவை. அகரவரிசையில் அமைந்த சொற்களஞ்சியங்களில் சிக்கல் என்னவென்றால் அவற்றில் தேவைக்கு அதிகமான பதிவுகள் காணப்பெறும். இரு சொற்கள் ஒருபொருள் பன்மொழிகள் என்றால் கூட அகரவரிசைக்காக இரண்டுமுறை நோக்கீடு செய்துதான் தேவையான தகவல்களைப் பெரும்பாலும் பெற இயலும். முதல்முறை அகரவரிசையில் அமைந்த சொல்லடைவு மூலம் குறிப்பிட்ட சொல் எப்பொருண்மைத் தலைப்பில் வருகிறது என்பதைத் தெரிந்துகொண்டு இரண்டாவது முறை நோக்கீடு செய்துதான் வேண்டிய தகவலைப் பெற இயலுகின்றது. இதனால் தேடும் நேரம் இரண்டு மடங்காக உயரும். ஆனால் கணிப்பொறியின் வழி தேடலை விரைவாகச் செய்ய இயலும். சொல் வலை மின்சொற்களஞ்சியத்தைவிட ஒரு படி மேலானது. தமிழ்ப் பல்கலைக்கழக மொழியியல் துறையில் இராசேந்திரனை முதன்மை ஆய்வாளராகக் கொண்டு இந்திய அரசின்

தகவல்தொடர்பு மற்றும் மின்னணு அமைச்சின் நிதி நல்கையின் கீழ் 2006-2009 என்ற காலகட்டத்தில் திராவிடச் சொல்வலை உருவாகத்தின் ஒருபகுதியாகத் தமிழ்ச் சொல்வலை உருவாக்கம் தொடங்கப்பட்டது முதல் கட்டத்தில் 15000 ஒருபொருள் பன்மொழியக் குழுங்களும் இரண்டாம் கட்டத்தில் 15000 ஒருபொருள் பன்மொழியக் குழுமங்களும் உருவாக்கப்பட்டு TDIL இணைய தளத்தில் பதிவேற்றம் செய்யப்பட்டுள்ளன.

சுருக்கமாக, சொற்களை அவற்றிற்கு இடையே காணப்படும் உறவுகள் அடிப்படையில் பிணைக்கும் ஒரு ஒழுங்குமுறை சொல்வலை (wordNet) என வரையறை விளக்கம் செய்யலாம். சொல்வலையின் அடிப்படை உருவாக்கக் கூறு ஒருபொருள்பன்மொழியக் குழுமங்களாகும் (synsets). ஒருபொருள்பன்மொழியக் குழுமங்கள்தாம் சொல் உறவுகளாலும் பொருண்மை உறவுகளாலும் தொடர்புபடுத்தப்பட்டுள்ளன. சொல்வலையின் பின்னணியில் மூலப்பொருண்மை ஆய்வு (ontology) அமையும்.

மொழி ஒரு ஒருங்கிணைந்த, ஒன்றிணைக்கப்பட்ட ஒழுங்குமுறையாகச் செயல்படுகின்றது. அடிப்படையில் மொழி உலகை மனித மூளையுடன் தொடர்புபடுத்துகின்றது. புலனால் உலகைக் கருத்துருவாக்கம் செய்ததன் விளைவாக மொழியின் அடிப்படை அலகுகளான சொற்கள் உருவாக்கப்பட்டன. அவை மூளைக்கு உள்ளீடுகளாக அமைந்து கருத்துப் பரிமாற்றத்தின் வெளியீடாக அமைந்தது. எனவே சொல்வலையை ஒரு மன/உள அகராதி (Mental dictionary) எனக் கூறலாம்.

சொல்வலையின் முக்கியப் பயன்பாடு சொற்கள் வழியாகத் தகவல்களைப் பெற உதவும் ஒழுங்குமுறையாக (Information Access system) அமைவதாகும். சொல்வலை பல பொருண்மை உறவுகளில் ஒரு கருத்துருவை மற்றொரு கருத்துருவுடன்

இணைப்பதன் காரணமாக அது இயல்பாகவே ஒரு எடுத்துக்காட்டான சொல் பெறும் ஒழுங்குமுறையாக (Word access system) மாறியுள்ளது. சொல்வலை பல நடைமுறைப் பயன்பாடுகள் கொண்ட ஒரு சொல் தரவு மையமாகும் (lexical database).

சொல்வலையின் பயன்பாடுகளில் ஒன்று மொழிபெயர்ப்பாகும். சொல்வலை ஒரு கருத்துருவை (concept) மற்றொரு கருத்துருவுடன் சொல் உறவுகளாலும் (lexical relations) பொருண்மை உறவுகளாலும் (semantic relations) தொடர்புபடுத்துவதோடு மட்டுமல்லாமல், ஒரு சொல்லின் சூழல் அடிப்படையிலான பொருண்மை வேறுபாடுகளையும் தருகின்றது. சொல்வலை பொருண்மையைச் சூழலுடன் தொடர்புபடுத்தும் தன்மை கொண்டிருப்பதன் காரணமாக அது ஒரு சொல்லின் பொருண்மைகளைச் சூழல் அடிப்படையில் வேறுபடுத்த இயலும்.

இந்தோசொல்வலை

சொல்வலைகள் உலகமுழுவதும் பல மொழிகளுக்காக உருவாக்கப்பட்டுள்ளன. எடுத்துக்காட்டாக ஆங்கிலம், செருமன், பிரஞ்சு, இத்தாலி போன்ற ஐரோப்பிய மொழிகளுக்குச் சொல்வலைகள் உருவாக்கப்பட்டுப் பயன்பாட்டில் உள்ளன. இவைகளை ஒன்றிணைக்கும் பன்மொழிய யூரோ சொல்வலையும் (Euro WordNet) உருவாக்கப்பட்டுள்ளது. யூரோ சொல்வலையைப் பின்பற்றி இந்திய மொழிகளுக்கும் இந்தோ சொல்வலை (Indowordnet) உருவாக்கும் செயல்பாடு தொடங்கப்பட்டுச் செயல்பட்டு வருகின்றது. இந்திக்கான சொல்வலை தான் இந்தியாவில் உருவாக்கப்பட்ட முதல் சொல்வலையாகும். தற்போது மராத்தி, குஜராத்தி, ஒரியா, பெங்காளி, அசாமீஸ், கொங்கணி, பஞ்சாபி, போடோ, மணிப்பூரி, சமஸ்கிருதம் ஆகிய மொழிகளுக்குச் சொல்வலைகள் உருவாக்கும் செயல்பாடு நடைபெற்றுவருகின்றது.

திராவிடச் சொல்வலை

இந்திய அரசின் மின்னணு மற்றும் தகவல் தொழில் நுட்ப நிறுவனத்தின் (Department of Electronics and Information Technology) நிதி நல்கையின் கீழ் திராவிட மொழிகளுக்கான சொல்வலைகள் உருவாக்கும் செயல்பாடும் தொடங்கப்பட்டு செயல்பட்டு வருகின்றது. தமிழ் (தமிழ்ப் பல்கலைக் கழகம்), மலையாளம் (அமிர்தா பல்கலைக் கழகம்ஆ), தெலுங்கு (திராவிடப் பல்கலைக் கழகம்), கன்னடம் (மைசூர் பல்கலைக் கழகம்) ஆகிய திராவிட மொழிகளுக்குச் சொல்வலைகள் உருவாக்கும் செயல்பாடு கடந்த சில ஆண்டுகளாக நடைபெற்றுவருகின்றது. திராவிட மொழிகள் இலக்கண அடிப்படையிலும் சொற்றொகுதிகள் அடிப்படையிலும் இந்தோ ஆரிய மொழிகளிலிருந்து வேறுபடுவதன் காரணமாகத் திராவிடச் சொல்வலை உருவாக்கம் முக்கியமானதாகக் கருதப்படுகின்றது. தற்போது இந்தி சொல்வலையை அடிப்படையாகக் கொண்டே பிற இந்திய மொழிகளுக்குச் சொல்வலை உருவாக்கும் செயல்பாடு நடைபெற்று வருகின்றது. பின்வருவன திராவிடச் சொல்வலை உருவாக்குவதன் நோக்கங்களாகும்

- திராவிடமொழிகளுக்கு மிக விரிவான, மிகச் சிறந்த பன்மொழிய பொருண்மைசார் சொல் தரவுமையம் உருவாக்குவது.
- மொழி வலையமைப்பைத் தொடர்புபடுத்தி மொழிச் சுதந்திரமான பொருண்மை கருத்துருக்களை உருவாக்குவது.
- எல்லாத் திராவிட மொழிகளுக்கும் பொருண்மைசார் தகவல்களை வகைப்படுத்துவதைத் தரப்படுத்துவது மற்றும் பயன்பாடுகளுக்கு மூலவளங்களை உருவாக்குவது.

யூரோ சொல்வலை

சொல்வலையைப்பற்றிக் கூறுகையில் யூரோ சொல்வலையைப் பற்றிக் கூறாது செல்ல இயலாது. ஏனென்றால் இது தான் முதன் முதலில் நடை முறைப்படுத்தப்பட்டு இணையத்தில் பயன்பாட்டில் உள்ளது. யூரோ சொல்வலையானது சொல் உறவுகளால் தொடர்புபடுத்தப்பட்ட சொற்களின் பன்மொழியச் சொல் வலையைத் (multi lingual wordnet) தருகின்றது. இது மொழிகளுக்கு இடையிலான தேடலைச் சாத்தியமாக்குகின்றது. சொல்வலையானது பெயர்கள், வினைகள், வினையடைகள், பெயரடைகள் இவை பற்றிய செய்திகளைக் கொண்டிருக்கின்றது. மேலும் இவை ஒருபொருள்பன்மொழியக் குழுமக் (synset) கருத்துருவைச் சுற்றி ஒழுங்கமைக்கப்பட்டுள்ளது. யூரோ வலையின் முக்கிய நோக்கம் என்னவென்றால் சிக்கனமான முறையில் பலமொழிகளுக்கிடையே (English, Dutch, Italian, Spanish, German, French, Estonian, Czech) மிக விரிவான மற்றும் தரம் வாய்ந்த பன்மொழியத் தரவு மையத்தை (Data base) உருவாக்குவதாகும். யூரோநெட்டினால் உருவாக்கப்பட்ட அறிவுவளம் பல தரப்பட்ட எல்லைக்குட்பட்ட பயன்பாட்டாளரால் பயன்படுத்தப்படுகின்றது.

தமிழ்ச் சொல்வலை

தமிழ்ச் சொல்வலையை (Rajendran 2009, 2010, 2016) மையமாகக்கொண்டு சொல்வலை உருவாக்குவது பற்றியும் அதன் பரிமாணங்கள் பற்றியும் இங்கு விளக்கப்பட்டுள்ளது.

அகராதியும் சொல் வலையும்

அகராதியானது சொல்லின் எழுத்துவடிவு, பொருள், பயன்பாடு இவற்றைக் காணப் பொதுவாய்ப் பயன்படுத்தப்படுகின்றது. ஆனால் சொல் வலை ஒரு பொருண்மை அகராதியாகும் (semantic lexicon). சொல்வலை சொற்களுக்கு வர்ணனையும் விளக்கமும் மாதிரி எடுத்துக்காட்டுகளும் தருமும். சொற்களுக்கு இடையே உள்ள பொருண்மை உறவுகளை மிகத்தெளிவாகப் பயன்படுத்தக்க விதத்தில் கூறும். உச்சரிப்பு, சொல்லாக்க உருபனியல் செய்திகள், சொல்மூலம் அல்லது வரலாறு இவற்றைத் தராது.

சொற்களஞ்சியமும் சொல் வலையும்

முதல் சொற்களஞ்சியம்/பொருட்புல அகராதி (Thesaurus) ராஜேஸ்ட் என்பரால் உருவாக்கப்பட்டது. இது பொருண்மை ஒழுங்கமைப்பைக் கொள்கையாகக் கொண்டு உருவாக்கப்பட்டுள்ளது. கருத்துரு அடிப்படையில் அமைக்கப்பட்டது. பயன்பாட்டாளர் தங்கள் மனதிலுள்ள கருத்துருவைக் கொண்டு சொற்களைக் கண்டுகொள்ள இது உதவும். சொல்வலையும் சொற்களஞ்சியம் போல அமையும். இதன் கட்டுமான அலகு ஒருபொருள்பன்மொழியக் குழுவும் ஆகும். ஒருபொருள்பன்மொழியக் குழுமத்தில் ஒரு கருத்துருவை வெளிப்படுத்தும் எல்லாச் சொற்களும் குழுமம் செய்யப்பட்டிருக்கும். சொல்வலையைப் பயன்படுத்துபவர் மனதிலுள்ள கருத்துருவைக் கொண்டு சொற்களைக் கண்டு கொள்ள இயலும். ஒருபொருள்பன்மொழியக் குழுமங்கள் உள்ளடங்கு-உள்ளடக்கு உறவு (honymy-hypernymy), சினை-முதல் உறவு (part-whole or meronymy-holonymy), உட்படுத்து உறவு (entailment) போன்ற உறவுகளால் இணைக்கப்பட்டுள்ளன. சொல்வலை கருத்து நிலையையும் சொல் நிலையையும் தெளிவாகப் பிரிக்கின்றது. இந்த வேறுபாடு சொல் உறவுகளுக்கும் கருத்துக்களுக்கும்

உள்ள வேறுபாட்டில் பிரதிபலிக்கின்றது. சொற்களஞ்சியத்திற்கு அப்பாற்பட்டு சொல்வலையில் சொற்களுக்கும் கருத்துருக்களுக்கும் உள்ள உறவுகள் வெளிப்படையாகக் கூறப்பட்டு அடையாளப்படுத்தப்பட்டுள்ளன. பயன்படுத்துபவவர் உளவுகளைத் தெரிந்தெடுத்து, ஒரு கருத்துருவிலிருந்து மற்றொரு கருத்துருவுக்குச் சென்று கருத்துரு வெளியிடத்தில் வலம் வர இயலும்.

சொல் வலையில் உறவுகள்

சொல்வலை கருத்துரு-பொருண்மை உறவுகளுக்கும் சொற்களை இணைக்கின்ற சொல் உறவுகளுக்கும் வேறுபாடு காட்டுகின்றது. ஆனால் சொல் நோக்கில் கவனம் செலுத்துபவர் முதன்மையாகச் சொல் உறவுகளைப் பயன்படுத்துவர். சொல்வலை பொருண்மை உறவுகளால் ஒழுங்கமைக்கப்பட்டுள்ளது. சொல் உறவுகள் ஒருபொருள்பன்மொழியக் குழுமங்களுக்கு இடையே உள்ள காட்டிகளாகும். சொல்வலை வேறுபட்ட தொடரியல் வகைப்பாடுகளில் அடங்கும் சொற்களை இணைக்கும் அடுக்கு உறவுகளைக் கொண்டிருக்கவில்லை. நான்கு முக்கியத் தொடரியல் வகைப்பாடுகளான பெயர், வினை, பெயரடை, வினையடை என்பன தனித்தனியாகத் தரப்பட்டுள்ளன. பெயர்கள் படிநிலை அமைப்புகளாகவும் வினைகள் வேறுபட்ட உட்படுத்து உறவுகளாலும் (entailment relation) பெயரடைகளும் வினையடைகளும் N-பரிமாண உயர்வெளியிடத்திலும் ஒழுங்கமைக்கப்பட்டுள்ளன.

பெயர்ச்சொல்வலை

பெயர்ச்சொல் வலையில் அடிப்படை பொருண்மை உறவு ஒருபொருள்பன்மொழிய உறவாகும். ஒருபொருள்பன்மொழியக் குழுமங்கள் அடிப்படைக் கட்டுமான அலகுகளாகும். ஒருபொருள்பன்மொழியக் கருத்துச்சாயல்

என்பது எல்லாச் சூழல்களிலும் ஒன்றையொன்று இடம்பெயர்த்தல் என்பதை உட்படுத்தாது. இடம்பெயர்த்தல் என்ற அளவியால் இயற்கை மொழிகளுக்குக் குறைந்த எண்ணிக்கையிலான ஒருபொருள்பன்மொழிகள் மட்டுமே வரும். சொல்வலையில் ஒருபொருள்பன்மொழிகளைக் குறைந்தது சில சூழல்களில் இடம் பெயர்க்கலாம். ஒருபொருள்பன்மொழியக் குழுமத்திற்கு ஒரு பொருள்தான் உண்டு.

எ.கா.

(நூல், புத்தகம், புஸ்தகம்) “படிப்பதற்கு ஏற்றவகையில் அட்டைப்போட்டு இணைத்த அச்சடித்த தாள்களின் தொகுப்பு”

சொல்வலையில் ஒருபொருள்பன்மொழியக் குழுமங்கள் பொருண்மை உறவுகளால் இணைக்கப்பட்டுள்ளன. பெயர்களை ஒழுங்குபடுத்தப் பயன்படுத்தப்படும் மிக முக்கியமான உறவு உள்ளடங்கு உறவாகும் (hyponymy). இந்த உறவுதான் சொற்களைச் சொல்படிநிலை அமைப்பில் ஒழுங்குபடுத்துகின்றது.

சொற்களின் படிநிலை அமைப்பு

உள்ளடங்கு உறவுகள் பற்றிய செய்திகள் மரபு அகராதிகளில் விவரணைகளாகத் தரப்பட்டிருக்கும். எ.கா.

குயில்	இனிய குரலுடைய கரிய நிறப்பறவை.
பறவை	இரு கால்களும் அலகுகளும் உடைய, உடலில் இருபக்கங்களிலும் பறப்பதற்கு ஏற்றவகையில் சிறகுகளும் கொண்ட விலங்கினம்
விலங்கினம்	தானாகவே இயங்கும் உணர்வு உறுப்புகளும் செல்லுலோஸ் இல்லாத செல் சுவர்களும் உள்ள உயிரினம்.

உயிரினம்	உயிர் வாழ்கின்ற ஒன்று.
----------	------------------------

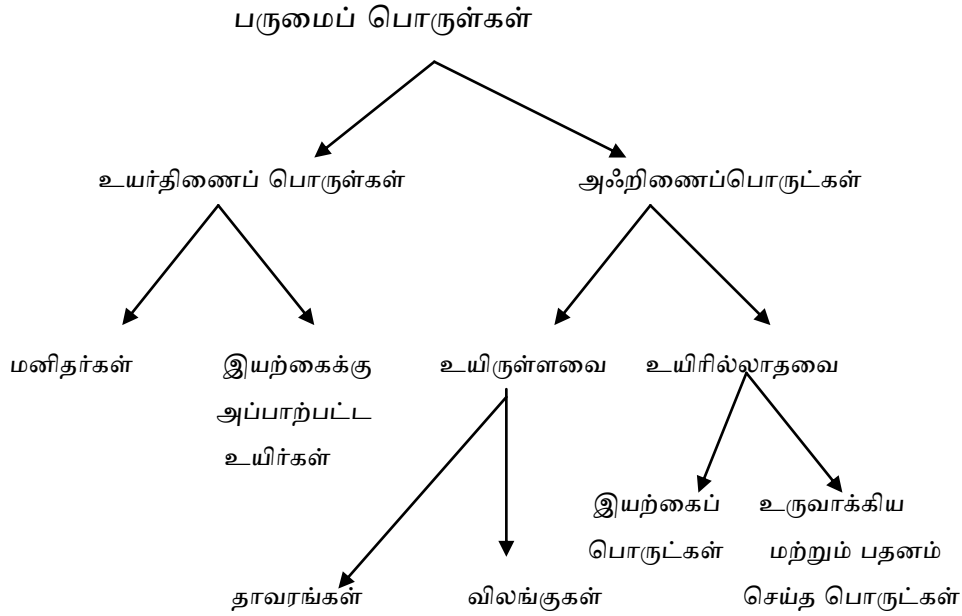
ஒவ்வொரு உள்ளடக்குச் சொல்லும் மிகப்பொதுவான உள்ளடக்குச் சொல்லுக்கு கொண்டு செல்லும். உள்ளடக்கு உறவைச் சொற்களுக்கு இடையே உள்ள உறவாக உருப்படத்தம் செய்ய இயலாது. உள்ளடக்கு உறவு அகராதிப்படுத்தப்பட்ட கருத்துக்களுக்கு இடையே உள்ள உறவாகும். ஒரு சொற்படிநிலை அமைப்பை உள்ளடக்கு உறவால் இணைக்கப்பட்ட ஒருபொருள்பன்மொழியக் குழுமங்களின் தொடர்ச்சியால் மீட்டுருவாக்கம் செய்யலாம்.

{குயில்}@->{பறவை} @ -> {விலங்கினம்}@ -> {உயிரினம்}

உள்ளடக்கு உறவுக்கு இணையான உள்ளடக்கு உறவையும் தரலாம். உள்ளடங்கு-உள்ளடக்கு உறவுகளால் ஒரு படிநிலை ஒழுங்குமுறை வெளிப்படும். இந்த வகையில் படும் படிநிலைகள் கணியியலார்களால் அறிவை உருப்படுத்தம் (knowledge representation) செய்யும் வழியாகப் பயன்படுத்தப்படுகின்றது.

தனித்தன்மையான தொடக்கிகள்

பெயர்கள் எல்லாவற்றையும் ஒரே படிநிலை அமைப்பில் தரவேண்டி படிநிலைக் கொள்கையை நீட்சி செய்யலாம். சொல்வலை பெயர்களைத் தனித்தன்மையான தொடக்கிகள் (unique beginners) அமையப் படிநிலை அமைப்புகளாகப் பகுத்துள்ளது. தனித்தன்மையான தொடக்கிகள் சொற்பொருண்மையியல் பொருண்மைக்கூறாய்வின் பொருண்மைக் கூறுகளுடன் ஓரளவுக்குப் பொருந்தும்.



மூலப்பொருண்மை அமைப்பைக் (ontological structure) காட்டும் வகையிலான சொற்களின் உருப்படுத்தம் சொற்களின் பொருண்மை மரபுரிமைச் செயற்பாங்கை (Lexical inheritance) வெளிக்கொணர்கின்றது.

எ.கா.

{ஊர்தி, வாகனம்}@ -> {நிலத்தில் ஓடும் ஊர்தி}@ -> {பொறியால் ஓடும் ஊர்தி}

@ -> நான்குசக்கர ஊர்தி}@ -> {பயண ஊர்தி} @ -> {பேருந்து}

யூரோ சொல்வலையின் பெயர் மூலக் கோப்புகளுக்கு 25 தனித்தன்மையான தொடக்கிகளின் பட்டியல் (List of 25 unique beginners for noun source files of Euro WordNet)

{act, activity}	{natural object}
{animal, fauna}	{natural phenomenon}
{artifact}	{person, human being}

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

{attribute}	{plant, flora}
{body}	{possession}
{cognition, knowledge}	{process}
{communication}	{quantity, amount}
{event, happening}	{relation}
{feeling, emotion}	{shape}
{food}	{state}
{group, grouping}	{substance}
{location}	{time}
{motivation, motive}	

வேறுபடுத்தும் பொருண்மைக் கூறுகள்

பெயர்களின் படிநிலை அமைப்பு உள்ளடங்கு-உள்ளடக்கு உறவால் உருவாக்கப்பட்டாலும் விளக்கங்கள் ஒரு கருத்துருவை மற்றொரு கருத்துருவிலிருந்து பிரிக்கும் பொருண்மைக் கூறுகளால் தரப்படுகின்றது.

எ.கா.

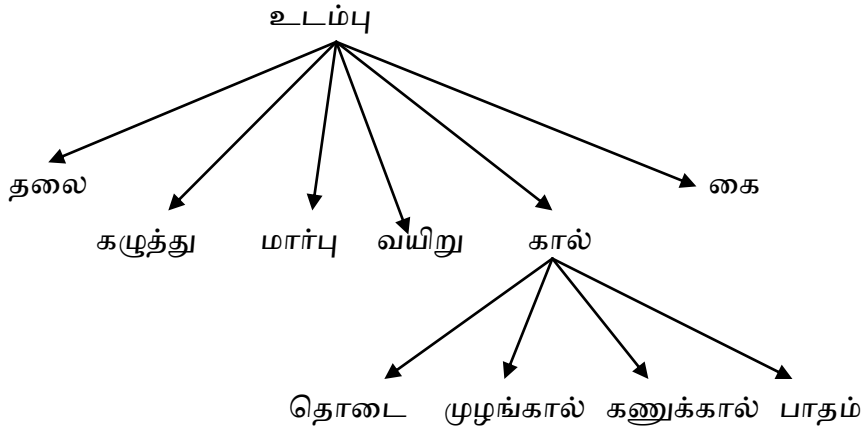
குயில் – ஒரு பாடுகின்ற சிறிய கரிய நிறப்பறவை

குயில் என்பதை மூன்றுவகைப் பட்ட வேறுபடுத்தும் பண்புக்கூறுகளுடன் தொடர்புபடுத்தலாம்.

1. அடைகள்: சிறிய, கரிய
2. பாகங்கள்: அலகு, சிறிய
3. செயல்பாடுகள்: பாடு, பற

சினை-முழு உறவு

சொற்களைப் படிநிலை அமைப்பில் தருவதில் சினை-முழு உறவு முக்கியப் பங்கு வகிக்கின்றது.



சினை-முழுஉறவு பின்வரும் துணை வகைகளை உள்ளடக்கும்:

1. முழுப் பொருளுக்கும் அதன் உறுப்புகளுக்கும் இடையிலான உறவை வெளிப்படுத்துவன (பாகம், எ.கா. கை - விரல்)
2. முழுமைக்கும் அதிலிருந்து பிரிந்த பகுதிக்கும் உள்ள உறவை வெளிப்படுத்துவன (பகுதி, எ.கா. துண்டு - உலோகம்)
3. இடங்களுக்கும் விரிந்த இடங்களுக்கும் உள்ள உறவை வெளிப்படுத்துவன (இடம், எ.கா. பாலைவனச்சோலை: பாலைவனம்)
4. குழுமத்திற்கும் அதன் அங்கத்தினர்களுக்கும் உள்ள உறவை வெளிப்படுத்துவன (எ.கா. மந்தை - ஆடு)
5. பொருள்களுக்கும் அது உருவாக்கப்பட்ட உருப்பொருள்களுக்கும் உள்ள உறவை வெளிப்படுத்துவன (ஆன, எ.கா. புத்தகம் - காகிதம்)

பெயர்ச் சொல்வலையின் பொருண்மை உறவுகளின் சுருக்கம்

பின்வரும் அட்டவணையில் பெயர் சொல்வலையில் கூறப்பட்டுள்ள

பொருண்மை-சொல் உறவுகள் பட்டியலிடப்பட்டுள்ளன:

உறவுகள்	துணை வகைகள்	எடுத்துக்காட்டுகள்
ஒருபொருள் பன்மொழி		புத்தகம், நூல்
உள்ளடங்கு-உள்ளடக்கு		விலங்கு-பலூட்டி
உள்ளடக்கு-உள்ளடங்கு		பசு-பாலூட்டி
முழு-சினை	முழுமை-பாகங்கள்	மேசை-கால்
“	பாகம்-முழுமை	சக்கரம்-வண்டி
“	அங்கத்தினர்கள்-குழு	படைத்தலைவர்-படை
“	குழு-அங்கத்தினர்கள்	துறை-பேராசிரியர்
“	பகுதி-முழுமை	துளி-கண்ணீர்
“	இடம்-பரந்த இடம்	பாலைவனச்சோலை- பாலைவனம்
“	பொருள்-உருப்பொருள்	புத்தகம்-தாள்
ஈரிணை எதிர்நிலை	நிரலாக்கம் செய்யத்தக்கது	நல்லவன் – கெட்டவன்
“	துணை எதிர்நிலை	பகல் – இரவு
“	தனிப்பட்டவை	அஃறிணை- உயர்திணை
“	துருவ எதிரிநிலை	ஆண் – பெண்

“	பரஸ்பர சமூகப்பாத்திரங்கள்	மருத்துவர் – நோயாளி
“	சொந்தங்கள்	அம்மா – மகள்
“	கால உறவுகள்	காலை – மாலை
“	இட உறவுகள்:செங்கோண	வடக்கு-கிழக்கு, மேற்கு
“	உறவுகள்	
“	இட உறவுகள்: நேர் எதிர்நிலை	வடக்கு –தெற்கு
“	பல்லிணை எதிர்நிலை: சங்கிலி	ஒன்று, இரண்டு, மூன்று
“	சுற்று	ஞாயிறு, திங்கள், செவ்வாய், புதன், வியாழன், வெள்ளி, சனி

வினைச் சொல் வலை

ஒரு குறிப்பிட்ட செயலை அல்லது தெரிவைக் குறிக்கும் சொல் வினைகள் ஒரு மொழியின் சொல் மற்றும் தொடரியல் வகைப்பாட்டில் முக்கிய பங்கு வகிக்கின்றது. இதன் பயனிலை – பங்கெடுப்பாளர் அமைப்பு (predicate argument structure) இது வரும் வாக்கியத்தின் சாத்தியமான அமைப்புகளைத் தீர்மானிக்கின்றது.

எ.கா.

காற்று வீசுகின்றது.

அவன் சென்னையிலிருந்து தஞ்சாவூர் வந்தான்.

அவள் மாம்பழம் சாப்பிடுகிறாள்.

அவர் அவளிடமிருந்து ஒரு லட்ச ரூபாய்க்கு ஒரு கார் வாங்கினார்.

வினைகளைச் சொற்களங்களாகப் பிரித்தல்

யூரோ சொல்வலை வினைகளை 15 சொற்களங்களாகப் பிரிக்கின்றது. தமிழ்ச் சொல்வலை வினைகளைத் தற்காலிகமாக (இராஜேந்திரன், 2001) பெளதிக வினைகள் (பெய், வீசு), உடல்கூறு வினைகள் (பிரசவி, வியர்), புலனுணர்வு வினைகள் (கசு, குளிர்), உணர்ச்சி வினைகள் (கோபப்படு, வேதனையடை), அறிவுசார் வினைகள் (ஊகி, கணி), கருத்துப்பாற்ற வினைகள் (பேசு, இகழ்), சலன வினைகள் (நகர், நட), தாக்க வினைகள் (அடி, உடை), உடைமைமாற்ற வினைகள் (கொடு, விடு), பல்கூட்ட வினைகள் (சமையல் செய்) எனப்பிரித்துள்ளது.

வினைகளுக்குத் தனித்தன்மையான தொடக்கிகள்

வினைச் சொற்களஞ்சியத்தைப் பொருண்மைக் களங்களாகப் பிரிப்பது வினைத் தரவுகளை ஒழுங்கு முறை படுத்துவதோடு சொற்களஞ்சியத்திலுள்ள எல்லா வினைகளுக்கும் ஒரு பொதுவான வேர் அல்லது தனித்தன்மையான தொடங்கிகள் இல்லாத குறையையும் நிவர்த்தி செய்கின்றது.

ஒருபொருள்பன்மொழியக்குழுமங்களாக வினைகள்

பெயர்களைப் போல வினைகளையும் ஒருபொருள் பன்மொழியக் குழுமங்களாகக் குழும இயலும். ஆனாலும் நாம் ஒருபொருள்பன்மொழி என்பது ஒன்றையொன்று இடம் பெயர்க்கும் சொற்கள் என்று வரையறை செய்தால் வினையில் ஒருபொருள்பன்மொழிகள் இல்லாமல் போகும் அல்லது மிக அரிதாகத் தான் இருக்கும்.

வினைகளின் பொருண்மைக் கூறாய்வு

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

வினைகளை அவற்றின் பொருண்மைப் பண்புக்கூறுகளால் விவரணை செய்ய இயலும். வினைகளைச் சிறு பொருண்மைக்கூறுகளாகப் பிரிக்க இயலும் தன்மை காரணமாக எளிய செயல்களால் கலவைத் தன்மையான செயல்களை விளக்க இயலும்.

கொல் - சாகச் செய்தல்

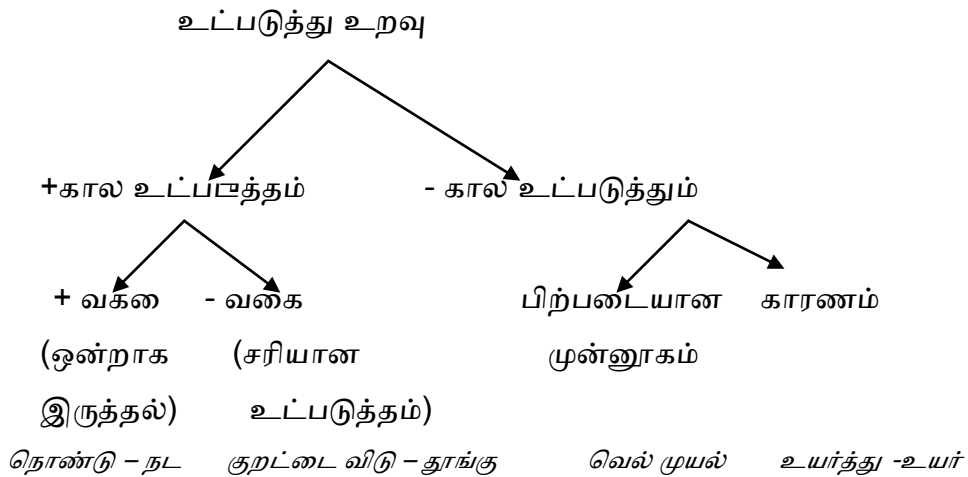
எறி - ஒருவர் தன் கையிலிருந்து பொருளை விசையால் போகச் செய்தல்

ஓட்டு - ஒருவர் ஒன்றை ஓடச் செய்தல்

வினைகளுக்கிடையிலான சொல் மற்றும் பொருண்மை உறவுகள்

நாம் நான்கு வகையான உட்படுத்து உறவுகளை (entailment)

வெளிப்படுத்தலாம்:



வினைகளுக்கிடையிலான பல்பொருள் ஒருமொழியம்

பெயர்களை விட வினைகளின் எண்ணிக்கை குறைந்தவை. வினைகளின் பல்பொருள் ஒருமொழியம் பெயர்களின் (polysemy) பல்பொருள் ஒருமொழியத்தைக் காட்டிலும் அதிகம். வினைகளின் பொருண்மை நெகழ்ச்சி அதன் பகுத்தாய்வைக் கடினப்படுத்துகின்றது. வினைகள் அவை எடுக்கும் பங்கெடுப்பவர் அமைப்பு

அடிப்படையில் (argument structure) தங்கள் பொருளை மாற்றும். ஆனால் பெயர்களின் பொருண்மை வினைகளின் பொருண்மைகளைக் காட்டிலும் நிரந்தரமானது.

தொடரியல் பண்புகளும் பொருண்மை உறவுகளும்

தற்போது அகராதியில் சொற்களுக்கு அதன் தொடரியல் பண்புக்கூறுகளைத் தரும் போக்கு காணப்படுகின்றது. வினைகளைப் பொருண்மை உறவுகளாக மட்டும் பார்ப்பதும் தொடரியல் பற்றிய சில செய்திகளைத் தரும். வினையைப்பற்றி தெரிந்து கொள்ள வினையின் தொடரியல் பண்புகளைச் சொல் வலையில் தருவது குறித்த சாத்தியம் ஆயப்பட வேண்டும்.

வினைச் சொல்வலையின் பொருண்மை உறவுகளின் சுருக்கம்

பின்வரும் அட்டவணையில் வினைச் சொல்வலையில் கூறப்பட்டுள்ள பொருண்மை-சொல் உறவுகள் பட்டியலிடப்பட்டுள்ளன:

உறவுகள்	விளக்கம்/துணை வகை	எடுத்துக்காட்டு
ஒருபொருள் பன்மொழியம்	இடம்பெயர்க்க இயலும் செயல்கள்	தூங்கு - உறங்கு
பகுதி - முழுமை	உட்படும் - உட்படுத்தம் செயல்	பற - பிரயாணி
வகை உறவு	செயல் - துணை வகை	நட - நொண்டு
உட்படுத்து உறவு	செயல் - காரணச்செயல்	உயர் - உயர்த்து
“	செயல் - முன்னூகச் செயல்	
எதிர்நிலை	எதிர்நிலை	கூடு - குறை
	மறுதலை	வில் - வாங்கு
	திசை எதிர்நிலை	புறப்படு - வந்துசேர்

பெயரடை மற்றும் வினையடைச் சொல் வலை

சொல் வலை பெயரடைகளை இரு முக்கியமான வகுப்புகளாகப் பிரிக்கின்றது: வர்ணனைப் பெயரடைகள் (descriptive adjectives) மற்றும் தொடர்புப் பெயரடைகள் (relational adjectives). வர்ணனைப் பெயரடைகள் தலைப் பெயர்களுக்கு இருதுருவ அடைகளின் மதிப்பீட்டைத் (values of bipolar attributes) தருகின்றது; எனவே ஈரிணை எதிர்நிலைகள் (binary oppositions) (எதிர்மறை-antonymy) தருகின்றது; மற்றும் பொருண்மை ஒற்றுமையால் (similarity of meaning) ஒழுங்கமைக்கப்படுகின்றது.

வர்ணனைப் பெயரடை (descriptive adjective): பெரிய, கனமான

தொடர்புப் பெயரடை (relational adjective): பொருளாதார, சகோதர

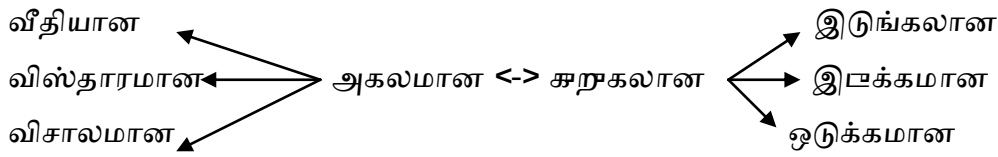
குறிப்பு அடைசெய்யும் பெயரடை (reference modifying adjectives): பழைய, முன்னாள்

ஒரு வர்ணனை அடை ஒரு பெயருக்கு ஒரு மதிப்பு தரும். எடுத்துக்காட்டாக, கனமான, இலேசான என்பது எடை என்பதன் அடைமொழிகளாகும்; தாழ்ந்த உயர்ந்த என்பன உயரம் என்பதன் அடைமொழிகளாகும். எதிர்மறை (antonymy) தான் வர்ணனைப் பெயரடைகளில் அடிப்படையானதாகும். எதிர்மறை () தான் வர்ணனைப் பெயரடைகளில் அடிப்படையானதாகும். எதிர்மறை ஒருபொருள்பன்மொழியத்தைப் போன்று சொல்வடிவுகளுக்கு இடையிலான உறவாகும். எதிர்மறைப் பெயரடைகள் அடைமொழியின் எதிர்செய்யும் மதிப்புகளை வெளிப்படுத்துகின்றது (எ.கா. கனமான, இலேசான என்பன எடை அடைமொழியின் எதிர் துருவங்களாகும்.) நேரடியான எதிர்மறைகள் இல்லாத பெயரடைகள் (எதிர்மறை ஆற்றல் குறைவாய் உள்ள பெயரடைகள்) எதிர்மறைப் பெயரடைகளுடன் “போல” உறவு கொண்டிருக்கும்:

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)



உறவுகள்	தொடர்புபடுத்தும் சொல்வகைப்பாடு	எடுத்துக்காட்டு
எதிர்மறை (நிரலாக்கம் செய்ய இயலுவது)	பெயரடை/பெயரடை	அழகான/குருமரான
எதிர்மறை (நிரலாக்கம் செய்ய இயலாதது)	பெயரடை/பெயரடை	உயிருள்ள/செத்த
ஆக்கவடிவு	பெயரடை/பெயர்	அழகான/அழகு
அடைமொழி	பெயர்/பெயரடை	வடிவம்/சின்ன
தொடர்பு	பெயரடை/பெயர்	பொருளாதார/பொருளாதாரம்
போல	பெயரடை/பெயரடை	பாரமான/கனமான

சொல்வலையை திட்டமிட்டு நடைமுறைப்படுத்தல்

சொல்வலையை நான்கு ஒழுங்கமைப்புகளாகப் பகுக்கலாம்.

1. சொல்வள ஒழுங்குமுறை (Lexical resource system)
2. ஒருங்கிணைக்கும் ஒழுங்கு முறை (Compiler system)
3. சேகரிப்பு ஒழுங்கு முறை (Storage system)
4. மீள்பெறும் ஒழுங்கு முறை (Retrieval system)

அகராதியியலார் சொல்வலையின் மூலக் கோப்புகளை எழுதுகின்றனர். இது சொல்வள ஒழுங்குமுறையை அமைக்கும். ஒருங்கிணைக்கும் ஒழுங்குமுறை சொல்வள கோப்புகளை தரவுமைய வடிவாக உருவாக்குகின்றது. சேகரிப்பு ஒழுங்குமுறை சொல்வள ஒழுங்குமுறைக்கும் ஒருங்கிணைப்பு ஒழுங்குமுறைக்கும் இடையில் செயலாற்றுகின்றது. மீளப்பெறும் ஒழுங்குமுறை பின்வரும் செயல்பாடுகளை உள்ளடக்கும்:

1. வகைப்பாட்டு மற்றும் வலைப்பின்னல் திட்டத்தை மின் தட்டச்சு அமைப்பாக மாற்றுகின்றது.
2. வகைப்பாட்டு மற்றும் வலைப்பின்னல் திட்டத்தை தகவல் மீட்பு ஒழுங்கமைப்பாக மாற்றுகின்றது.
3. இணைய மீளப்பெறுதல்: வகைப்பாட்டு மற்றும் வலைப்பின்னல் திட்டத்தை குறிதேடலுடன் இணைப்பது.

முடிவுரை

சொல்வலை ஒரு சொல் தரவு மையம் (Lexical database). இதன் முக்கியமான பண்பு அதன் அர்த்தங்களின் வலைப்பின்னல். சொற்களுக்கு இடையே உள்ள ஒருப்பொருண்மையை அளந்து சொற்பொருள் மயக்கம் தீர்க்க சொல்வலை பயன்படுகின்றது. சொல்வலையைப் பயன்படுத்தி படிநிலை அமைப்பில் தரப்படுள்ள ஒரு சொற்களுக்கு இடையே உள்ள பொருண்மை உறவை ஆய்ந்து அச்சொற்களுக்கு இடையே உள்ள ஒற்றுமையை அளக்கலாம். பொருண்மை மயக்கம் தவிர மீட்புச்செயல்பாடுகளுக்கும் சொல்வலை பயன்படுகின்றது. தகவல் தளங்களிலும் தகவல் மையங்களிலும் இருந்து பயன்படுத்துபவரின் கேள்வியால் தேவையான

தகவல்களைப் பெற இயலும். சொல்வலை இதற்குப் பெரிதும் உதவும். இயற்கை மொழி ஆய்வுக்கும் உருவாக்கத்திற்கும் இது பல வழிகளில் கைகொடுக்கும். சொல்வலையைப் பயன்படுத்தி இயந்திரமொழிபெயர்ப்பு செய்வதற்கான முயற்சிகளும் நடைபெற்று வருகின்றன. சொல்வலையில் உட்படுத்தப்பட்டுள்ள சொல் அறிவு அதைத் திறமையான ஒழுங்குமுறைகள் (expert systems), மொழிபெயர்ப்புத் துணைக்கருவிகள், தேடல் இயந்திரங்கள், கற்றல் ஒழுங்கமைப்புகள் மற்றும் தானியங்கு சுருக்கிகள் (automatic summarizers) போன்றவற்றிற்குப் பொருத்தமானதாகச் செய்கின்றது. யூரோ சொல்வலை போலவே திராவிடமொழிகளின் சொல்வலை (Dravidian wordNet) உருவாக்குவதற்கான முயற்சிகள் தொடங்கப்பட்டு நடைமுறைப் படுத்தப்பட்டு வருகின்றன.

நோக்கீட்டு நூல்கள் மற்றும் கட்டுரைகள்

இராசேந்திரன், ச. 1999. பொருட்புல வகைப்பாடும் சொற்களஞ்சியமும் [Classification into semantic domains and Thesaurus]. புலமை தொகுதி 25, எண்2, டிசம்பர், 1999, 47-66..

இராசேந்திரன், ச. 2001. தற்காலத் தமிழ்ச் சொற்களஞ்சியம். தமிழ்ப் பல்கலைக்கழகம். தஞ்சாவூர். இராசேந்திரன், ச. மற்றும் பாஸ்கரன், ச. 2006. தமிழ் மின்சொற்களஞ்சியம். தமிழ்ப் பல்கலைக்கழகம், தஞ்சாவூர்.

இராசேந்திரன், ச. 2002. "மொழி ஆய்வில் கணிப்பொறியின் பயன்பாடு [Use of computer in language analysis]". In: சித்திரபுத்திரன் மற்றும் பிறர் (பதிப்பு). ஆய்வியல் கட்டுரைகள். அன்னைஅகரம் பதிப்பகம். தஞ்சாவூர், 2002.

இராசேந்திரன், ச. 2006. தமிழ்ச் சொல்வலை. [Tamil WordNet]. In: ஆய்வுமுகங்கள். எச். சித்திரபுத்திரன் மற்றும் பிறர் (பதிப்பு). ஆசிரியர் பேரவை. தமிழ்ப் பல்கலைக்கழகம். தஞ்சாவூர், 2006, 105-117.

இராசேந்திரன், ச. மற்றும் ச. பாஸ்கரன். தற்கால மின்சொற்களஞ்சியம். தமிழ்ப் பல்கலைக்கழகம். தஞ்சாவூர்,

சிவசுப்பிரமணியக் கவிராயர் (மூலப்பதிப்பு1930, மறுபதிப்பு 1985). நாமதீபநிகண்டு. (உரையாசிரியர்: வையாபுரிப்பிள்ளை). தஞ்சாவூர்: தமிழ்பல்கலைக்கழகம்.

திவாகரர். 1958. சேந்தன் திவாகரம். சென்னை: கழகம்சுடி

ராமகிருஷ்ணன், எஸ். (பதிப்பாசிரியர்). க்ரியாவின் தற்காலத் தமிழ் அகராதி. சென்னை: க்ரியா.

Abel, A. (2012). Dictionary writing systems and beyond. In S. Granger & M. Paquot (Eds.),

Electronic lexicography (pp. 83-105). Oxford: Oxford University Press.

Abel, A., & Klosa, A. (2012). Der lexikographische Arbeitsplatz – Theorie und Praxis.

In R. V. Fjeld & J. M. Torjusen (Eds.), Proceedings of the 15th EURALEX International Congress (pp. 413-421). Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo.

Amsler, Robert A. 1980. Ph.D. Dissertation, "The Structure of the Merriam-Webster Pocket Dictionary". The University of Texas at Austin.

Atkins, B. T. S., & Rundell, M. (2008). The Oxford Guide to Practical Lexicography. Oxford: Oxford University Press.

Byrd, Roy J., Nicoletta Calzolari, Martin S. Chodorow, Judith L. Klavans, Mary S. Neff, and Omneya A. Rizk. "Tools and methods for computational lexicology." *Computational Linguistics* 13, no. 3-4 (1987): 219-240.

Cruse, D.A. 1986. *Lexical Semantics*. New York: Cambridge University Press

Cruse, D.A. 2000. *Meaning in Language: An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.

Guckler, G. 1983. Appendix: B: A Computer-based Monolingual Dictionary: A Case Study. In R.R.K. Hartmann (ed.) 1983. *Lexicography: Principles and Practice*. London: Academic Press Inc.

Hanks, P. (2012). The corpus revolution in lexicography. *International Journal of Lexicography*, 25, 398–436.

Jones, K.S. 1986. *Synonymy and Semantic Classification*. Edinburgh: Edinburgh University Press.

Karunakaran, K. 1987. Language Planning in Tamil: retrospects and prospects. *Tamil Civilization*, vol 5, no. 3, 58-65.

Kilgarriff, A. (2005). If dictionaries are free, who will buy them? *Kernerman Dictionary News*, 13, 17-19.

Lehman, Thomas & Martin, Thomas. 2007 edition. *A word Index of Cangam Literature*. Institute of Asian Studies.

Lew, R. (2010). Multimodal lexicography: The representation of meaning in electronic dictionaries. *Lexikos*, 20, 290-306.

Lew, R. (2012). How can we make electronic dictionaries more effective? In S. Granger & M. Paquot (Eds.), *Electronic lexicography* (pp. 343-361). Oxford: Oxford University Press.

Lyons, J. 1963. *Structural semantics*. Oxford: Blackwell.

Lyones, J. 1977. *Semantics* (vol.) Cambridge: Cambridge University Press.

Marello, Carla (1998) "Hornby's bilingualized dictionaries", in *International Journal of Lexicography* 11,4, pp. 292–314.

Martin, W.J.R., B.P.F.Al and P.J.G.Van Sterkenburg. 1983. On Processing of A text Corpus. In R.R.K. Hartmann (ed.). 1983. Lexicography: Principles and Practice. London: Academic Press Inc.

Mawson, C.O.S. 1956. Roget's International Thesaurus of English Words and Phrases. New York: Pocket Books, INC.

Nida, E.A. 1975a. Compositional Analysis of Meaning: An Introduction to Semantic Structure. The Hague: Mouton.

Nida, E.A.. 1975.b. Exploring Semantic Structure. The Hague: Mouton

Nielsen, Sandro (1994) The Bilingual LSP Dictionary, G. Narr. ISBN 978-3-8233-4533-6

Nielsen, Sandro (2008) "The effect of lexicographical information costs on dictionary making and use", in Lexikos (AFRILEX-reeks/series 18), pp. 170–189.

Nielsen, Sandro (2009): "Reviewing printed and electronic dictionaries: A theoretical and practical framework". In S. Nielsen/S. Tarp (eds): Lexicography in the 21st Century. In honour of Henning Bergenholtz. Amsterdam/Philadelphia: John Benjamins, 23–41. ISBN 978-90-272-2336-4

Rajendran, S. 1978. Syntax and Semantics of Tamil Verbs. Ph.D. Thesis. Poona: University of Poona.

Rajendran, S. 1983. Semantics of Tamil Vocabulary. (Report of the UGC sponsored Postdoctoral Work in manuscript). Poona: Deccan College Post Graduate and Research Institute.

Rajendran, S. 1995. "Towards a Compilation of a Thesaurus for Modern Tamil". South Asian Language Review. 5.1:62-99

Rajendran, S. 1996-2000. "A Model for a 'Theoretical Dictionary of Tamil". Tamil Civilization vol. 14-18, March 1996-2000, 99-105.

Rajendran, S. 1996. "The Feasibility of Preparing a Thesaurus using Corpus". Workshop on Indian Language Corpus and its Applications (28, 29 Oct. 1996). Central Institute of Indian Languages, Mysore.

Rajendran, S. 2002. Preliminaries to the preparation of Wordnet for Tamil. Language in India www.languageinindia.com 2:1, March 2002.

Rajendran, S. and Baskaran, S. 2002. Electronic Thesaurus for Tamil. In: Proceedings of the International Conference on Natural Language Processing. NCST, Mumbai, 2002. Rajendran, s. 2003. Prerequisite for the Preparation of an Electronic Thesaurus for a Text Processor in Indian Languages. Language in India www.languageinindia.com 3:1, January 2003.

Rajendran, S. 2009. Dravidian WordNet. In: Proceedings of Tamil Internet Conference 2009. Cologne, Germany, October, 2009

Rajendran, S. 2010. "Tamil WordNet." In: Proceedings of the Global WordNet Conference (GWC 10) 2010, IIT, Bombay.

Rajendran, S., Shivapratap G, Dhanalakshmi V and Soman K.P. 2010. "Building a WordNet for Dravidian Languages." In: Proceedings of the Global WordNet Conference (GWC 10), 2010, IIT Bombay.

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

Rajendran, S. Anand Kumar M. and Soman K.P. 2015. "Building hierarchies and networks from MRDs of Tamil". In: Proceedings of 14th International Conference on Tamil Computing and Tamil Internet (Tamil Internet 2015), at Singapore (SIM University, Singapore campus) from 30th, 31st May to 1st June 2015.

Rajendran, S. 2016. "Tamil Thesaurus to Tamil wordNet." In: Proceedings of 15th World Tamil Internet conference 2016, held at Gandhi Gram Rural University, Tamil Nadu, September 8-11, 2016, pages 1-9.

Rajendran, S. and Anandkumar, M. 2017. "Visual Onto-Thesaurus for Tamil." Language in India www.languageinindia.com ISSN 1930-2940 vol. 17:5 May 2017.

Rajendran S. 2018 Lexical Resource Tools for Tamil Computing: History, Development and Future Prospects. In: மாநாட்டுக் கட்டுரைகள், 17-ஆவது உலகத் தமிழ் இணைய மாநாடு (தமிழ்நாடு விவசாயப் பல்கலைக்கழகம், கோயம்பத்தூர், சூலை 6-8, 2018), உலகத்தமிழ் தகவல் தொழில் நுட்ப மன்றம் (உத்தமம்) (ISSN-2313-4887, California, USA. பக்கம் 14-15.

Rajendran S. 2018. "Visual Onto-thesaurus in Tamil." In: Jayadevan V. and Oppila Mathivanan (ed.) The Frontiers of Lexicography, Arimaa Nokku, Chennai pages 101-113.

Rundell, M. (2012). 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. In R. V. Fjeld & J. M. Torjusen (Eds.), Proceedings of the 15th EURALEX International Congress (pp. 47-92). Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo.

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

Slaughter, M.M. 1982. *Universal Languages and Scientific Taxonomy in the Seventeenth Century*. Cambridge: Cambridge University Press.

Trap-Jensen, L. (2010). One, two, many: Customization and user profiles in internet dictionaries. In A. Dykstra & T. Schoonheim (Eds.), *Proceedings of the XIV Euralex International Congress* (pp. 1133-1143). Ljouwert: Afûk.

Vaiyapuri Pillai, S. 1935. 'Introduction', in *Tamil Lexicon* (vol 1). Madras: Madras University Press.

Verlinde, S., Leroyer, P., & Binon, J. (2010). Search and you will find. From stand-alone lexicographic tools to user driven task and problem-oriented multifunctional leximats. *International Journal of Lexicography*, 23(1), 1-17. doi: 10.1093/ijl/ecp029.

Wikipedia: Electronic dictionary. Downloaded on 20/3/2023

Wikipedia: Lexicography. Downloaded on 20/3/2023

Wikipedia: List of online dictionaries,

வினாவங்கி

1. பொருத்தமான விடையைத் தேர்ந்தெடுத்தல் (15 வினாக்கள்)

1) சொல்லாட்சி அகராதி

அ. சொற்களை அகரவரிசைப் படுத்தித்தரும்

ஆ. சொற்கள் நூலில் இடம்பெறும் இடத்தைத் தரும்

இ) எத்தனைமுறை, எந்தச் சூழலில் பயன்படுத்தப்படுகின்றது என்பனவற்றைத் தரும்

ஈ) மேற்சொன்ன மூன்றையும் தரும்

2) கணினிசார் சொல்லியியல் ஒரு தனி ஒழுக்கமாக

Language in India www.languageinindia.com ISSN 1930-2940 **23:3 March 2023**

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

அ. 1960களில் வெளிப்பட்டது√

ஆ. 1970களில் வெளிப்பட்டது

இ. 1980களில் வெளிப்பட்டது

ஈ. 1990களில் வெளிப்பட்டது

3) ஃப்ரேம் நெட்டை உருவாக்கி வளத்தவர்

அ. பைக் வாசன்

ஆ. பெல்பாம்

இ. ஃப்மோர்

ஈ. மில்லர்

4) சொல்வலை முதல்முதலில்

அ. பிரஞ்சு மொழியில் உருவாக்கப்பட்டது

ஆ. ஸ்பானிஷ் மொழியில் உருவாக்கப்பட்டது

இ. இத்தாலிய மொழியில் உருவாக்கப்பட்டது

ஈ. ஆங்கில மொழியில் உருவாக்கப்பட்டது√

5) அகராதியைப் பயன்படுத்துவதன் காரணம்

அ. புதிய சொல்லின் அர்த்தத்தை அறிந்துகொள்ள

ஆ. புதிய சொல்லின் உச்சரிப்பை அறிந்துகொள்ள

இ. புதிய சொல்லின் எழுத்துக்கூட்டலை அறிந்துகொள்ள

ஈ. மேற்சொன்ன மூன்றையும் அறிந்துகொள்ள √

5) முதல் முதலில் சொல்லடைவை கணிப்பொறிவழி உருவாக்கம் செய்தவர்

அ. மாத்தூஸ் அர்னால்ட்√

ஆ. ஸ்டிபன்

இ. ஃபில்மோர்

ஈ. ஃபெல்பாம்

6) அகராதி தயாரிப்பதில் முதற்கட்டப் படிநிலை

அ. சொல்லுக்கான இலக்கணக் கூறுகளை ஆய்தல்

ஆ. பொருள் எழுதுதல்

இ. பயன்பாட்டை விளக்குதல்

ஈ. செய்திகளைச் சேகரித்தல், ஆராய்ந்து வரிசைப்படுத்தி அகரவரிசைப் படுத்தல்

7) சொற்களஞ்சியம் என்பதில்

அ. சொற்கள் அகரவரிசையில் தரப்பட்டிருக்கும்

ஆ. சொற்களின் ஒருபொருள்பன்மொழிகள் மட்டும் தரப்பட்டிருக்கும்

இ. சொற்களின் எதிர்மொழிகள் மட்டும் தரப்பட்டிருக்கும்

ஈ. சொற்களின் எல்லாப் பொருண் உறவுகளும் வெளிப்படும்படி சொற்கள் தரப்பட்டிருக்கும்

8) இராசேந்திரனின் சொற்களஞ்சியத் தரவுத்தளம்

அ. நைடாவின் பொருண்மைக்கூறாய்வு ஆய்வு அடிப்படையில் உருவாக்கப்பட்டது

ஆ. ஜாண்லியானின் பொருண்மையியல் கோட்பாடுகள் அடிப்படையில் உருவாக்கப்பட்டது

இ. குருசின் பொருண்மையியல் கோட்பாடுகள் அடிப்படையில் உருவாக்கப்பட்டது

ஈ. காட்சின் பொருண்மையியல் கோட்பாடுகள் அடிப்படையில் உருவாக்கப்பட்டது.

9) சொல்வலை என்பது

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

அ. ஒரு அகராதி

ஆ. ஒரு சொற்களஞ்சியம்

இ. அகராதியும் சொற்களஞ்சியமும் இணந்த ஒரு சொல்சார் தரவுத்தளம்

ஈ. ஒரு இயந்திரம் படிக்கவியலும் அகராதி

10) நூல், புத்தகம் இவைகளுக்கு இடையே உள்ள பொருண்மை உறவு

அ. உள்ளடங்கு-உள்ளடக்கு மொழியம்

ஆ. சினை-முழுமை மொழியம்

இ. ஒருபொருள் பன்மொழியம்

ஈ. பல்பொருள் ஒருமொழியம்

12) நாற்காலி, கால் இவைகளுக்கு இடையே உள்ள பொருண்மை உறவு

அ. உள்ளடங்கு-உள்ளடக்குமொழியம்

ஆ. சினை-முழுமை மொழியம்

இ. ஒருபொருள் பன்மொழியம்

ஈ. எதிர்மொழியம்

13) காட்சிச்சொற்களஞ்சியத்தில் ஒருபொருள்பன்மொழிகள் தரப்பட்டுள்ள

அடைப்புக்குறிகள்

அ. {}

ஆ. < > √

இ. / |

ஈ. \$ #

14) பொருளாதார, சகோதர என்பன

அ. வர்ணனைப் பெயரடைகளாகும்

ஆ. குறிப்பு அடைசெய்யும் பெயரடைகளாகும்

இ. விளக்கப் பெயரடைகளாகும்

ஈ. தொடர்புப் பெயரடைகளாகும்\

15) குறட்டைவிடு, உறங்கு என்பவைகளுக்கு உள்ள உறவு

அ. உள்ளடங்கு-உள்ளடக்கு மொழிய உறவாகும்

ஆ. பகுதி-முழுமை மொழிய உறவாகும்

இ. உட்படுத்து உறவாகும்

ஈ. மறுதலை ஆகும்

2. பொருத்துக (10 வினாக்கள்)

1) விளக்கங்களைச் சொல்வள மொழிக்கருவிகளுடன் பொருத்துக

அ. சொற்களை அகரவரிசையில் தருவது – சொற்களஞ்சியம்

ஆ) சொற்களை அகரவரிசைப் படுத்தியும் நூலில் இடம்பெறும் இடங்களையும்

பயன்படுத்தப்படும் சூழலையும் தருவது – சொற்பொருளடைவு

இ) மேற்சொன்னவற்றுடன் பொருள் அடிப்படையில் பிரித்துத்தருவது - சொல்லாட்சி

அகராதி

ஈ) ஒருபொருள்பன்மொழியம் போன்ற சொல்லுறவுகள் அடிப்படையில் சொற்களைக்

குழுமித்தருவது – சொல்லடைவு

2) எடுத்துக்காட்டுகளைப் பொருண்மை உறவுகளுடன் பொருத்துக

அ) நூல், புத்தகம் – எதிர்மொழிகள்

ஆ) யானை, விலங்கு – சினை-முழு மொழிகள்

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

இ) தலை, உடல் - உள்ளடங்கு-உள்ளடக்கு மொழிகள்

ஈ) மேலே, கீழே - ஒருபொருள் பன்மொழிகள்

3) காட்சிச் சொற்களஞ்சியத்தில் பயன்படுத்தப்படும் சொற்குழுமிகளைப் பொருத்துக

அ) இணைமொழியம் - % *

ஆ) வகைமொழியம் - / |

இ. பகுதிமொழியம் - { }

ஈ. எதிர்மொழியம் - < >

4) பொருண்மை உறவுகளை எடுத்துக்காட்டுகளுடன் பொருத்துக

அ. எதிர்மொழிகள் - கிளி, பறவை

ஆ. ஒருபொருள் பன்மொழிகள் - கை, விரல்

இ. பகுதி-முழுமொழியம் - நேரம், காலம்

ஈ. உள்ளடங்கு-உள்ளடக்கு மொழியம் - முன்னர், பின்னர்

4) சொற்களஞ்சியன்களை தொகுத்தவருடன்/எழுதியவருடன் பொருத்துக

அ) ஆங்கில மொழி சொற்கள் மற்றும் சொற்றொடர்களின் சொற்களஞ்சியம் - எட்வின்

எல். பாட்டிஸ்டெல்லா

ஆ) அமெரிக்க மொழியின் அமெரிக்கன் ஹெரிடேஜ் டிக்ஷனரி - ஆர்.எஸ். பி. பீக்ஸ்

இ) கிரேக்கத்தின் சொற்பிறப்பியல் அகராதி - க்டன் மிஃப்ளின் ஹர்கார்ட்

ஈ) சொற்களஞ்சியம் ஜாக்கிரதை - பீட்டர் மார்க் ரோஜெட்

5) தொகுத்தவரை சொற்களஞ்சியத்துடன்/நூலுடன் பொருத்துக

அ. பெ. மாதையன் மற்றும் எச். சித்திரபுத்திரன் - சங்க இலக்கியத்தில் விலங்கின

விளக்கம்

ஆ. ச. இராசேந்திரன் – நாபதீபநிகண்டு

இ. சிவசுப்பிரமணியக் கவிராயர் – தற்காலத் தமிழ்ச் சொற்களஞ்சியம்

இ. பி.எல்.சாமி – மர இனப்பெயர் தொகுதி

6) நூலை எழுதியவருடன் பொருத்துக.

அ. அகராதியியல் – ஆர்.எ. சிங்

ஆ. தமிழ் அகராதியியல் - லாடிஸ்லவ் சுசுஸ்தா

இ. அகராதியியல் கையேடு – எச் சித்திரபுத்திரன்

ஈ. அகராதியியல் அறிமுகம் – பெ. மாதையன்

7) ஆசிரியரை நூலுடன் பொருத்துக

அ. எச் சித்திரபுத்திரன் - இருபதாம் நூற்றாண்டுத் தமிழ்ப்பேரகராதி

ஆ. இரா. லியோ நிர்மலா – Onomatopoeia in Tamil

இ. வி. ஞானசுந்தரம் – தமிழ் இரட்டிப்புச் சொற்களின் அமைப்பு வகைப்பாடும் சொற்பொருள் தனித்தன்மையும்

ஈ. பி. இராமநாதன் – தமிழ் இரட்டிப்புச் சொற்களின் அமைப்பும் சொற்பொருண்மையும்

8) உறவுமுறைச் சொற்களைப் பொருத்துக

அ. ஒப்பன் - அண்ணன் மனைவி

ஆ. காக்கா – அம்மா

இ. தொத்தா – அண்ணன்

ஈ. மாட்டுப்பெண் – அப்பா

9) நூல்களை ஆசிரியர்களுடன் பொருத்துக

அ. தமிழ்ச்சொல்லகராதி – கு. சண்முகம்பிள்ளை

Language in India www.languageinindia.com ISSN 1930-2940 23:3 March 2023

Prof. S. Rajendran

Developing Lexical Resourceful Aids Based on Computational Technology (Tamil Textbook)

ஆ. தமிழ் தமிழ் அகரமுதலி – மு. சண்முகம்பிள்ளை

இ. தமிழ் மொழியகராதி – நா. கதிர்வேற்பிள்ளை

ஈ. தற்காலத் தமிழ் சொல்லகராதி – ச. பாவானந்தம்பிள்ளை

10) ஆசிரியர்களை நூற்களுடன் பொருத்துக:

அ. வ. ஜெயதேவன் - கிளைமொழியியல்

ஆ. அ.கா. பெருமாள் - வட்டார வழக்குச் சொல்லகராதி

இ. கி. ராஜநாராயணன் - நாஞ்சி வட்டார வழக்குச் சொல்லகராதி

ஈ. கோ. சீனிவாச வர்மா - தமிழ் அகராதியியல் அன்றும் இன்றும்

3. சரியா தவறா (15 வினாக்கள்)

1) dictionary என்ற சொல் dictionarium என்ற லத்தீன் சொல்லிலிருந்து பெறப்பட்டது.

√சரி/தவறு

2) ஃப்ரேம்நெட் என்பதை உருவாக்கியவர் பைக்வாசன்.

சரி/தவறு√

3)1987ஆம் ஆண்டில் பைர்ட், கால்சோலாரி, சோடோரோ என்போர்

உரைப்பகுப்பாய்விற்கான கணினிசார் கருவிகளை உருவாக்கியுள்ளனர்.

√சரி/தவறு

4) ஒருபொருள் பன்மொழிகள் மேலுறல் நிலையில் வரும்

சரி/தவறு√

5) மாத்தூஸ் அர்னால என்பவர் சொல்லடைவை முதன்முதலாக கணிப்பொறிவழி

உருவாக்கினார்.

√சரி/தவறு

6) நூல், மேஜை என்பன ஒருபொருள் பன்மொழிகள்.

சரி/தவறு√

7) தமிழ்ப்பலகளைக்கழகத் தற்காலத் தமிழ்சொற்களஞ்சியத்தைத் தொகுத்தவர் இராசேந்திரன் அவர்கள்.

√சரி/தவறு

8) கால்-உடல் என்பன உள்ளங்கு-உள்ளடக்கு மொழியத்திற்கு எத்துக்காட்டுகள்.

சரி/தவறு√

9) சக்கரம், வண்டி என்பன சினை-முழு மொழியத்திற்கு எடுத்துக்காட்டுகள்.

√சரி/தவறு

10) நட, நொண்டு என்பன எதிர்மொழிய உறவை வெளிப்படுத்தும்.

சரி/தவறு√

11) வில், வாங்கு என்பன மறுதலை உறவை வெளிப்படுத்தும்.

√சரி/தவறு

12) நட, நொண்டு என்பன திசை எதிர்நிலைகளாகும்.

சரி/தவறு√

13) உயர், உயர்த்து என்பன செயல்-காரணச்செயல் என்பதை வெளிப்படுத்தும்.

√சரி/தவறு

14) ஆழகான, குருரமான என்பன நிரலாக்கம் செய்ய இயலா எதிர்மொழிகளாகும்.

சரி/தவறு√

15) அழகான என்பது அழகு என்பத ஆக்கவடிவம் ஆகும்.

√சரி/தவறு

4. ஒரு சொல்/சொற்றொடரில் விடைதருக. (15 வினாக்கள்)

- 1) சொற்களை அகரவரிசைப்படுத்தித் தருவது
- 2) சொற்களை அகரவரிசைப்படுத்தி வரும் இடம் சூழல் போன்ற தகவல்களைத்தருவது.
- 3) ஒருபடி மேல் சென்று பொருள் அடிப்படையில் பிரித்துத் தருவது.
- 5) கணினிசார் அகராதியியல் என்றால் என்ன?
- 6) கணினிசார் சொல்லியியல் என்றால் என்ன?
- 7) முதல் முதலில் உரைப்பகுப்பாய்விற்கான கணினிசார் கருவிகளை உருவாக்கியவர்கள்.
- 8) பிரேம் நெட் ஒர்க்கை உருவாக்கி வளத்தவர் யார்.
- 9) தமிழ்சொல்வலை முதன்முதலில் உருவாக்கப் பட்ட பல்கலைக்கழகம்.
- 10) ஒருபொருள் பன்மொழியம் என்றால் என்ன?
- 11) உள்ளடங்கு உள்ளடக்கு மொழியம் என்றால் என்ன?
- 12) சினை-முழு மொழியம் என்றால் என்ன?
- 13) திசை எதிர்மொழியத்திற்கு இரு எடுத்துக்காட்டுகள் தருக.
- 14) நிகண்டு என்றால் என்ன?
- 15) தமிழ் அகராதியியல் அன்றும் இன்றும் என்ற நூலை எழுதியவர் யார்?

5. ஒருபத்தியில் விடை தருக. (10 வினாக்கள்)

- 1) சொல்லாட்சி அகராதி என்றால் என்ன?
- 2) இருமொழிய அகராதி என்றால் என்ன? எடுத்துக்காட்டுகளுடன் விளக்கு.
- 3) சொல்லியலுக்கும் அகராதியியலுக்கும் உள்ள வேறுபாடு என்ன?
- 4) அகராதி பயன்படுத்துவதற்கான காரணங்களைச்சுருக்கமாகக் கூறுக.

5) சொற்களஞ்சியம் என்றால் என்ன?

6) தற்கப்பொருண்மை உறவுகளைச் சுருக்கமாகக் கூறுக.

7) ஒருபொருள் பன்மொழியத்திற்கும் பல்பொருள் ஒருமொழியத்திற்கும் இடையிலுள்ள வேறுபாட்டை எடுத்துக்காட்டுடன் சுருக்கமாகக் கூறுக.

8) இந்தோ சொல்வலை என்றால் என்ன?

9) சொல்வலைக்கும் மின்சொற்களஞ்சியத்திற்கும் உள்ள வேறுபாடு என்ன?

10) சொற்றொகையின் படிநிலை அமைப்பு என்றால் என்ன?

6. மூன்றுபக்க அளவில் விடைதருக. (5 வினாக்கள்)

1) கண்ப்பொறி வழி அகராதி உருவாக்குவது பற்றி ஒரு கட்டுரை வரைக.

2) கண்ப்பொறி வழி சொற்களஞ்சியம் உருவாக்குவது பற்றி எடுத்துக்காட்டுகளுடன் விளக்குக.

3) மொழிக்கருவிகளை உருவாக்கும்போது எதிகொள்ளப்படும் சிக்கல்களையும் தீர்வுகளையும் விளக்குக.