கம்பராமாயணத்தில் சொல்வள வளர்ச்சி வீதம்

Relative Growth of vocabulary (RGV)
in *kambarAmAyanam*

1. Relative Growth of vocabulary (RGV) **என்றால் என்ன**?

vocabulary **என்பதைச்** சொல்வளம் எனக் குறிப்பிடலாம். ஒரு குறிப்பிட்ட பகுதியில் (text) உள்ள தனிச்சொற்களை அதன் vocabulary எனலாம். அதாவது, அப் பகுதியில் 1000 சொற்கள் இருந்தால், பல சொற்கள் ஒன்றுக்கும் மேற்பட்ட தடவைகள் பயன்படுத்தப்பட்டிருக்கும். அவ்வாறு அதிகப்படியாக வரும் சொற்களைத் தவிர்த்துவிட்டு ஒவ்வொரு சொல்லையும் ஒருமுறைமட்டுமே எடுத்துக்கொண்டால் அதுவே அப்பகுதியின் சொல்வளம் ஆகும். 1000 சொற்கள் உள்ள பகுதியில் 250 தனிச்சொற்கள் இருப்பின், அப்பகுதியின் சொல்வளம் 250/1000 = 25/100 = 25% எனலாம். எனவே அதன் சொல்வளம் குறைவு எனக் கொள்ளலாம். அதாவது மிகச் சிலவான சொற்களே அங்கு திரும்பத் திரும்பப் பயன்படுத்தப்பட்டிருக்கின்றன என்பது அதன் பொருள். மாறாக, அங்கு 800 தனிச்சொற்கள் இருப்பின் அதன் சொல்வளம் 80% எனலாம். அப்படிப்பட்ட பகுதியைச் சொல்வளம் மிக்க பகுதி எனலாம். அதாவது, அங்கே நிறைய புதுச்சொற்கள் பயன்படுத்தப்பட்டிருக்கின்றன எனலாம். அப்படிப்பட்ட பகுதி வாசிப்போரின் சொல்வளத்தைக் கூட்டும்.

இப்போது, ஒரு பெரிய பகுதியில் மொத்தம் 1000 சொற்கள் இருப்பதாகக் கொள்வோம். இதனை 10 சமபாகங்களாகப் பிரிப்போம். இப்போது முதல் 100 சொற்களில் 45 தனிச் சொற்கள் இருப்பதாகக் கொள்வோம். இது இப் பகுதியில் 45/100 = 45% ஆகிறது. தொடர்ந்து அடுத்த 100 சொற்களை எடுத்து, இப்போது முதல் 100+100=200 சொற்களில் சுமார் 90 புதிய தனிச்சொற்கள் இருக்கும் என்று சொல்லமுடியாது. ஏற்கனவே முதல் 100 சொற்களில் இருக்கும் தனிச்சொற்களில் பல இங்கு மீண்டும் பயன்படுத்தப்பட்டிருக்கும். அவைபோக, இப்போது முற்றிலும் புதிய தனிச்சொற்கள் 30 இருக்கலாம். எனவே, முதல் 200 சொற்களில் இப்போது 45+30=75 தனிச் சொற்கள் உள்ளன. இது <u>மொத்தத்தில்</u> 75/200 = 38% ஆகிறது. எனவே, தனிச்சொற்கள் கூடுதலாகக் கிடைத்தாலும் அவற்றின் வளர்ச்சி குறைந்துகொண்டுவருவதைக் காணலாம். தொடர்ந்து, அடுத்த 100 சொற்களையும் எடுத்து, இப்போது முதல் 300 சொற்களைக் கொண்டால், இப்போது தனிச்சொற்கள் இன்னும் குறைவாக இருக்கும். இப்போது 300 சொற்களில் தனிச்சொற்கள் 100 எனக்கொண்டால் இது <u>மொத்தத்தில்</u> 100/300 = 33% ஆகிறது. எனவே, தனிச்சொற்களின் விழுக்காடு மேலும் குறைந்திருக்கிறது. அதாவது தனிச்சொற்களின் எண்ணிக்கை மேலும் மேலும் கூடினாலும், அவற்றின் வளர்ச்சிவீதம், கொஞ்சம் கொஞ்சமாகக் குறைந்துகொண்டே வருவதைக் காணலாம். இதைத்தான் Relative Growth of vocabulary (RGV) என்கிறோம். இதைப்போல் ஒவ்வொரு ஆயிரம் சொற்களையும் எடுத்து, தொடர்ந்து கணக்கிட்டால், போகப் போக, தனிச்சொற்களின் எண்ணிக்கை குறைந்துகொண்டே வருவதைக் காணலாம். ஒரு ஆசிரியர் எழுதிக்கொண்டே போகும்போது, எவ்வளவு

நேரம்தான் அவருக்குப் புதிய தனிச்சொற்கள் தோன்றிக்கொண்டே இருக்கும்? இது அவரின் சொல்வளத்தைப் பொருத்தது. எப்படி இருந்தாலும், ஒரு மிகப் பெரிய பகுதியை எடுத்துக்கொண்டால், ஒரு அளவுக்குப் பிறகு அவரின் சொல்வளம் குறைந்துகொண்டே சென்று, அவரின் RGV 0-வை எட்டும். இனிமேல் அவர் ஏற்கனவே எழுதிய சொற்களைத்தான் பயன்படுத்திக்கொண்டே இருப்பார். எனவே இருவேறு ஆசிரியர்களின் ஒரே அளவு பகுதியை எடுத்து, அப்பகுதியில் அவர்களின் RGV - ஐக் கண்டால் அவர்களின் சொல்வளத்தை ஒப்பிடமுடியும். இதை, மனித முயற்சியில் செய்வது மிகவும் கடினம். எனவே, இதற்கான கணினி நிரல்கள் ஆசிரியரால் எழுதப்பட்டு, கம்பராமாயணத்தின் அனைத்துப் பாடல்களுக்கும் RGV கண்டுபிடிக்கப்பட்டுள்ளது.

கொள்கை அளவில் மேற்சொன்ன எடுத்துக்காட்டு சரியாகப்பட்டாலும், பயன்பாட்டில் மிகவும் கவனமாக இதனைக் கையாளவேண்டும். முதலில் சொல் என்றால் என்ன என்பதில் தெளிவு தேவை. இந்த ஆய்வுக்காக இந்த ஆசிரியரால் ஒரு சொற்பிரிப்புக் கொள்கை உருவாக்கப்பட்டு, அதன் முறையில் சொற்கள் பிரிக்கப்பட்டு கணினி நிரல்களால் கணக்கிடப்பட்டுள்ளன. அடுத்து, தனிச்சொற்கள் எவை என்பதிலும் கொள்கைகள் வேறுபடுகின்றன. படி = study , படி = step படி = a measure என்பவை ஒரே சொல்லா அல்லது வெவ்வேறா? படிக்க, படித்து, படித்தான், படித்தாள் ஆகியவை வெவ்வேறு சொற்களா அல்லது ஒரே சொல்லின் இலக்கண வேறுபாடுகள்தானா (grammatical variations)? இன்னோரன்ன பல இடற்பாடுகள் உண்டு. கணினி நிரல்கள் சொற்களின் உருவ வேறுபாடுகளை வைத்தே சொற்களைக் கணக்கிடுகின்றன. இரண்டு சொற்களில் எழுத்துக்கள் ஒரே மாதிரி இருந்தால் அவை ஒரே சொல்லாகவும், ஏதேனும் ஓர் எழுத்து மாறுபட்டால் அவை தனித்தனிச் சொற்களாகவும் கணினி கொள்ளும்.

மேலும் ஒரு ஆசிரியர் எதற்காக எழுதுகிறார்?; யாருக்காக எழுதுகிறார்?; எதைப்பற்றி எழுதுகிறார்?; என்பதைப் பொருத்து அவரின் நடையும், சொல்தேர்வும் வேறுபடலாம். இவற்றையும் கணக்கில் கொள்ளவேண்டும். சிறுவர்க்கான ஒரு நூலில், ஆசிரியர் தன் சொல்வளத்தை எல்லாம் காட்டியிருப்பார் எனக் கொள்ளமுடியாது. ஓர் இலக்கிய இதழுக்கும், ஒரு சாதாரண (வெகுஜன) இதழுக்கும் கதை எழுதும் ஒரே ஆசிரியர், இரண்டிலும் ஒரே சொல்வளத்தைக் காட்டுவார் என எதிர்பார்க்க முடியாது. எனவே, ஒப்பீட்டுப் பின்புலங்கள் மிகவும் ஒத்துப்போகும் பகுதிகளையே சொல்வள ஒப்பீட்டுக்கு எடுத்துக்கொள்ள வேண்டும்.

2. கம்பராமாயணத்தில் RGV

இந்த ஆய்வுக்காக எடுத்துக்கொண்ட சொற்பிரிப்பு நெறிகளின்படி, கம்பராமாயணத்தில் மொத்தம் 2,40,568 சொற்கள் இருக்கின்றன. இதனைப் பத்துப்பாகங்களாகப் பிரிக்கிறோம். ஒவ்வொரு பாகமும் 24,056 சொற்கள் கொண்டிருக்கும். இந்த ஒவ்வொரு பாகத்திலுமுள்ள 24,056 சொற்களில் ஏற்கனவே பயன்படுத்தப்பட்ட சொற்களைத் தவிர புதிய சொற்கள் எத்தனை என்று கணக்கிடவேண்டும். முதல் 24,056 சொற்களில் 9575 தனிச் சொற்கள் கிடைக்கின்றன. இது 39.8% ஆகும். அடுத்த 24,056 சொற்களையும் சேர்த்து மொத்தம் 48,112 சொற்களில் 15,635 தனிச் சொற்கள் கிடைக்கின்றன. பார்த்தீர்களா! இப்போது நமக்கு 15,635 - 9575 = 6060 புதிய தனிச் சொற்கள்தான் கிடைத்துள்ளன.

அடுத்த 24,056 சொற்களையும் சேர்க்க, இப்போது 48,112 + 24,056 = 72,168 சொற்களில் 20,877 தனிச் சொற்கள் ஆக இப்போது நமக்குக் கிடைத்திருக்கும் புதுச் சொற்கள் 20,877 - 15635 = 5242 மட்டுமே. ஆக நாம் ஒவ்வொரு 24,056 சொற்களாக அதிகரித்துக்கொண்டே செல்லும்போது நமக்குக் கிடைக்கும் புதிய தனிச் சொற்களின் எண்ணிக்கை 9575, 6060, 5242 எனக் குறைந்துகொண்டே வருகிறது. எனவே, அதிகரித்துக்கொண்டுவரும் தனிச் சொற்களின் வளர்ச்சிவீதம் குறைந்துகொண்டே வருவதைக் காண்கிறோம். ஒரு குறிப்பிட்ட நிலைக்குப் பின்னர், இந்தப் புதிய தனிச் சொற்களின் எண்ணிக்கை மிகவும் குறைந்து, தனிச் சொற்களின் எண்ணிக்கை ஏறக்குறைய ஒரு நிலையான எண்ணிக்கையை அடையும் என எதிர்பார்க்கலாம். ஆனால் கம்பனின் இராமாயணத்தைப் பொருத்தமட்டில் அவ்வாறு நடைபெறவில்லை என்பதனைக் காணலாம்.

இந்த அட்டவணையைக் கவனியுங்கள்.

எண்.	சொற்களின் எண்ணிக்கை	விழுக்காடு	தனிச்சொற்கள்(%)		புதிய தனிச்சொற்கள்	%
1.	24,056	10%	9575	20.75%	9575	39.8%
2.	48,112	20%	15,635	33.88%	6060	25.2%
3.	72,168	30%	20,877	45.24%	5242	21.8%
4.	96,224	40%	25,565	55.40%	4688	19.5%
5	1,20,280	50%	29,693	64.34%	4128	17.2%
6.	1,44,336	60%	33467	72.52%	3774	15.7%
7.	1,68,392	70%	36732	79.60%	3265	13.6%
8.	1,92,448	80%	39987	86.65%	3255	13.5%
9.	2,16,504	90%	43065	93.32%	3078	12.8%
10	2,40,560	100%	46,144	100%	3079	12.8%

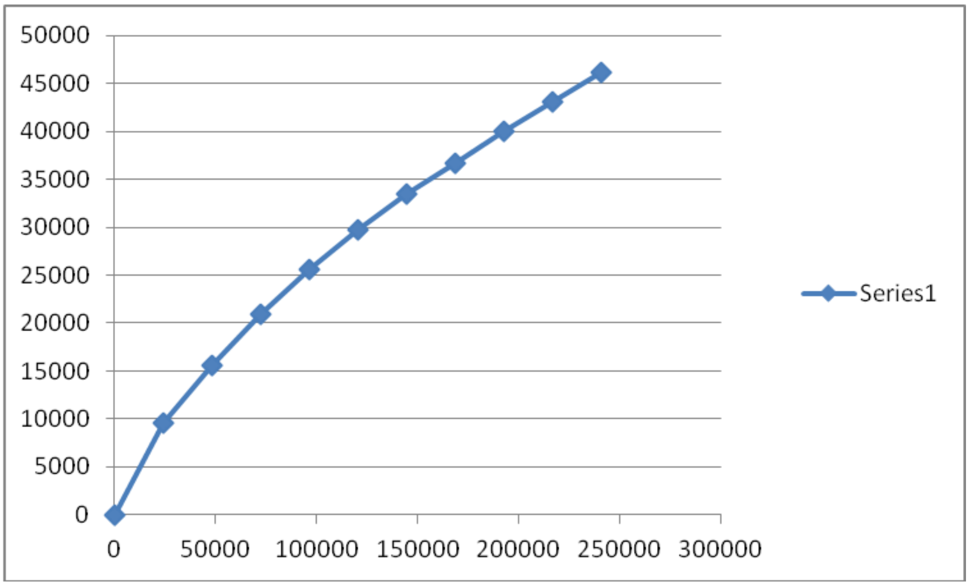
கம்பராமாயணத்தில் உள்ள மொத்தச் சொற்கள் 2,40,568 எனவும் இதில் உள்ள தனிச்சொற்கள் 46,144 எனவும் பார்த்தோம்.

இங்குள்ள அட்டவணைகளில் 7 நிரல்கள் (Colums)உள்ளன. நிரல் 1. வரிசை எண். கம்பராமாயணத்தில் உள்ள மொத்தச் சொற்கள் 10 சமபாகங்களாகப் பிரிக்கப்பட்டுள்ளன. அதாவது கம்பனில் உள்ள மொத்தச் சொற்களின் எண்ணிக்கை மட்டமாக்கப்பட்டு (round off) 2,40,560 ஆகிறது. இதனைப் பத்துப்பத்தாகப் பிரிக்க, 24,056, 48,112, 72,168, 96,224….. என்ற எண்கள் பெறப்படுகின்றன. இவை நிரல்-2 இல் தரப்பட்டுள்ளன. மொத்தச் சொற்களில் இவை 10%, 20%, 30%, 40% என்றாகிறது இல்லையா! இந்த விழுக்காடு நிரல் - 3 இல் கொடுக்கப்பட்டுள்ளது. இந்த எண்ணிக்கைக்குள் அமைந்த தனிச்சொற்களைக் காணவேண்டும். அதாவது முதல் 24,056 சொற்கள் எடுக்கப்பட்டு, அவற்றில் உள்ள தனிச்சொற்கள் கணக்கிடப்படுகின்றன. அது 9575 ஆகும். இது நிரல் 4-இல் தரப்பட்டுள்ளது. இந்த 9575 சொற்கள், மொத்தத் தனிச்சொற்களான 46,144 இல் 20.75% ஆகும். இது நிரல்-5 இல் தரப்பட்டுள்ளது. இதே போல், அடுத்த 24,056 சொற்களையும் சேர்த்து மொத்தம் 48,112 சொற்களில் உள்ள தனிச்சொற்கள் 15,635. இதன் விழுக்காடு 33.88 எனப் பெறப்படுகிறது. இவ்வாறாக 10 நிரைகளும் கணக்கிடப்படுகின்றன.

எடுக்கப்பட்டுள்ள பத்துப் பகுப்புகளில் ஒவ்வொரு பகுப்பிலுமுள்ள 24,056 சொற்களில் கிடைக்கும் தனிச் சொற்கள் நிரல் - 6 இல் கொடுக்கப்பட்டுள்ளன.

நிரல் 4-இல் உள்ள எண்களில், ஒவ்வொரு எண்ணையும் அதன் முந்தைய எண்ணைக் கழிக்கக் கிடைக்கும் எண்கள் இவை. ஒவ்வொரு நிலையிலும் ஏற்கனவே பயன்படுத்தப்பட்டுள்ள தனிச் சொற்களைத் தவிர்த்துப் புதிதாக அந்தப் பகுப்பில் பயன்படுத்தப்பட்டுள்ள புதிய தனிச் சொற்களின் எண்ணிக்கையை இவை குறிக்கும். இவற்றின் எண்ணிக்கை போகப்போகக் குறைந்துகொண்டே வருவதைக் காணலாம். நிரை (Row) 5-ஐக் கவனியுங்கள். இது வரிசை எண் 5 -க்கு நேரானது. அதுவரையிலான மொத்தச் சொற்கள் 1,20,280. இதற்கு நேரான புதிய தனிச்சொற்களைப் பாருங்கள். அது 4128. இது அந்தப் பகுப்புக்குரிய 24,056 சொற்களில் 17.2% ஆகும். அதாவது சுமார் ஒரு லட்சத்துக்கும் அதிகமான சொற்களைப் பயன்படுத்திய பிறகும், அடுத்த பகுப்பில் நூற்றுக்குப் பதினேழு சொற்கள் புதிய சொற்கள் பயன்படுத்தப்பட்டிருக்கின்றன. இது கம்பரின் சொல்வளத்தைக் குறிக்கிறது இல்லையா! இதேபோல் கடைசி நிரையைக் கவனியுங்கள். காப்பியத்தைப் பாடி முடிக்கின்ற நிலையிலும், அந்த இறுதிப்பகுப்பில் 12.8% புதிய தனிச் சொற்களைப் பயன்படுத்தியிருக்கிறார் புலவர். ஒருவேளை இன்னும் பாடியிருந்தாலும், அதாவது இரண்டரை இலட்சம் சொற்களைப் பயன்படுத்திய பின்னரும், அதுவரை பயன்படுத்தியிராத புதிய தனிச் சொற்களை நூற்றுக்குப் பன்னிரண்டு என்ற விழுக்காட்டில் அவர் பயன்படுத்தியிருக்கிறார் என்பது பெரும் வியப்பைத் தருகிறது.

இதனை ஒரு வரைபடம் மூலவும் வரைந்து காட்டலாம். X-அச்சில் மொத்தச் சொற்கள் எண்ணிக்கையையும், Y-அச்சில் தனிச்சொற்கள் எண்ணிக்கையையும் எடுத்துக்கொண்டு, நிரல் 2-இல் உள்ள எண்களை X- அச்சிலும், நிரல் 4-இல் உள்ளவற்றை அவற்றுக்கான Y-அச்சு வழியும் குறித்து அந்தப் புள்ளிகளை இணைத்தால் கிடைக்கும் வரைகோடு சொல்வள வளர்ச்சியைக் குறிக்கும்.



ஆசிரியர் பயன்படுத்தும் ஒவ்வொரு சொல்லுமே தனிச்சொல்லாக இருந்தால் (கொள்கையளவில்) இந்த வரைகோடு ஒரு நேர்கோடாகத் தோன்றும். இதனோடு ஒப்பிடும்போது சொல்வள வரைகோடு முதலில் உயரே எழும்பி, போகப்போக உயரம் குறைந்துவருவதைக் காணலாம். புதுச்சொல் பயன்பாடு போகப்போக்க் குறைந்தால் இந்த வரைகோடு போகப் போக மிகவும் வளைந்து X-அச்சுக்கு இணையான ஒரு கோடாக மாறிவிடும். அதாவது மிக அதிகமான

சொற்கள் உள்ள ஒரு பெரிய பகுதியில், RGV (தனிச்சொல் வளர்ச்சி வீதம்) குறைந்துகொண்டே வந்து மேலும் வளர இயலாத நிலையை எட்டும் என உணரலாம். மாறாக்க் கம்பனில் இந்த வளைகோடு மேலும் மேலும் உயர்ந்துகொண்டே செல்வதைக் காண்கிறோம்.

கம்பராமாயணத்தில், இராமன் போருக்குப்புறப்படத் தன் வில்லையும், அம்புகள் நிறைந்த அம்பறாத்தூணியையும் எடுத்துக்கொண்டான் என்று சொல்ல வந்த கம்பர், இராமனது அம்பறாத்துணி, எடுக்க எடுக்க அம்புகள் குறையாததுஎன்று குறிப்பிடுகிறார். அதனைத் தொலைவு இலாத் தூணி என்கிறார் கம்பர். இதற்கு ஓர் அழகிய உவமையைக் கூறுகிறார் கம்பர். கவிஞர்களின் நாவிலிருந்து வரும் சொற்கள் போல இந்த அம்புகள் குறையாமல் வந்துகொண்டே இருக்கும் என்று குறிப்படும் கம்பர் கூறுகிறார்:

நல் இயல் நவை அறு கவிஞர் நா வரும்

சொல் எனத் தொலைவிலாத் தூணி தூக்கினான் - யுத்தகாண்டம்-முதற்போர்புரிபடலம் - 114.

இராமனின் இந்தத் தொலைவிலாத் தூணிபோன்றதே நம் கம்பரின் வற்றாத ஊற்றாய் பெருகிக்கொண்டு வரும் அவரின் சொற்கள் என்னும் அம்பாறாத்தூணி.