

எதிர்கால இணையத்தில் தமிழ் யூனிகோடு: வளர்தமிழ் ஆய்வில் தரவுத்தளங்களும், திரட்டிகளும்

முனைவர் நா. கணேசன், ஹ்யூஸ்டன், அமெரிக்கா

1.0 கணினி வலையில் தமிழ்

இந்திய மொழிகளிலேயே முதன்முதல் கம்ப்யூட்டரிலும், பின்னர் இணையத்திலும் ஏறிய மொழி தமிழ். தமிழ் எழுத்தின் எளிமை மற்ற இந்திய எழுத்துக்களை ஒப்பிட்டால் விளங்கும். ஏனை இந்திய எழுத்துப் போலன்றி, விராமப் புள்ளி தமிழில் கூட்டெழுத்துக்களை உடைத்தெறிகிறது. அச்சு, தட்டச்சு, கணியச்சு, இணையக் குழுக்கள், ... தமிழுக்கே முதல்-வரவான இவை புள்ளியின் நன்கொடை! கணிப்பொறிகளில் தமிழ் எழுத்துக்கள் 1980-களின் ஆரம்பத்திலே தெரிந்தன. அப்போது ஆப்பிள், மைக்ரோசாப்ட் கணினிகளில் பல எழுத்துருக்களும், எழுதிகளும் (editors), ஆர்வலர்கள் தோன்றவைத்தனர். ஜார்ஜ் ஹார்ட் (பெர்க்கிலி), ஆதமி (கே. சீனிவாசன்), மயிலை (கல்யாண்), நளினம் (செல்லையா), கம்பன் (வாசுதேவன்), அணங்கு (குப்புசாமி), துணைவன் (ரவி), அஞ்சல் (முத்தெழிலன்) ... போன்ற எழுதுரு, செயலிகளைக் குறிப்பிடலாம். இவை யாவும் இலவசமாகப் பயனர்களுக்குக் கிட்டியதால் பலப்பலரும் தத்தம் கணினிகளில் பொருத்தித் தமிழைப் பார்த்தனர். இருப்பினும், ஒவ்வொன்றும் தட்டெழுத்தத் தனியான எழுதிகளும் செயலிகளும் வேண்டியிருந்தன. வடிவான எழுத்துருக்கள் வேண்டித் தமிழ்நாட்டில் பத்திரிகை நிறுவனங்கள் தனித்தனி எழுதுருவாக வடித்துக் கொண்டனர். பத்திரிகை, புத்தகப் பதிப்பகங்களில் கணிஅச்சிடல் வேருன்றியது. முதன்முதலாக தினமலர் பத்திரிகை முழுமையும் கணிஅச்சில் இயங்கத் தொடங்கியது. பின்னர் கணினிகளால் படிப்படியாக ஈயஅச்சுக் கோர்ப்பதே ஒழிந்து இன்றைய நிலைக்கு உயர்ந்தது.

இணையம் உதயமாகி வேகமாக வளரத் தொடங்கியது 1994 முதல் என்று சொல்லலாம். அதற்குமுன் அமெரிக்காவின் பெரிய பல்கலைக் கழகங்களிலும், ஆய்வுக் கூடங்களிலும் மட்டுமே இணையம் புழங்கியது. இந்தியாவில் இணையம் 1990-களின் மத்தியில் பரவ ஆரம்பித்தது. இணையத்தின் வருகையால் தமிழ்க் கணிமை புதிய விசை பெற்றது, வைய விரிவு வலைப்பக்கங்கள் பிறந்தன. மின்னஞ்சல், அரட்டை, குழுமங்கள், ... என விரிந்தது. தமிழ்நெட் குழுவைப் பாலாபிள்ளை தொடங்கி நடத்தினார். 8-பிட் குறியீடு ஆகிய திஸ்கி உருவாகி யாகு குழுக்கள் தொடங்கப்பட்டன.

தொன்னூறுகளின் கடைசியில் இருந்து ஆறாம்திணை, இன்தாம், திண்ணை, அம்பலம், திசைகள், ... ஆதி மின்னிதழ்கள் குறிப்பிடத்தக்கன. பின் வணிக இதழ்கள் - குமுதம், விகடன், ... - இணையத்தில் வலம்வரத் தொடங்கின. விசைப்பலகைகள், எழுத்துக் குறியீடுகள் பலவாகிக் குழப்பங்களும் மிகுதியாக இருந்த சூழல் என்றும் குறிப்பிட்டாக வேண்டும்.

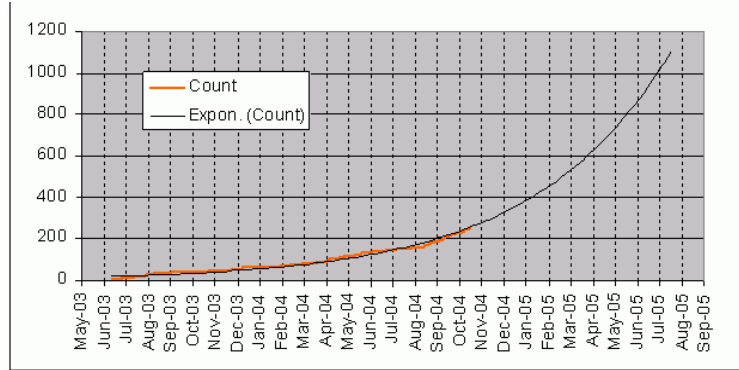
2.0 இணைய மாநாடுகளும், பல்கலை தரவுத்தளங்களும்

சிங்கப்பூரில் நா. கோவிந்தசாமி போன்றோர் ஒருங்கிணைந்து 1997-ல் முதல் இணைய மாநாடு நடத்தினர். பின்னர் சென்னை மாநாட்டில் தரக்கட்டுப்பாடாக தமிழ்99 விசைப்பலகை, டாப், டாம் குறியீடுகள் அரசாங்கம் அறிவித்தது. ரவீந்திரன் பால் வடித்த துணைவன் விசைப்பலகை தமிழ்99 பலகையின் முன்மாதிரி ஆகும். துணைவன் விசைப்பலகை இந்திய மொழிகளுக்கு முதல் ஒலியியல் (phonetic) விசைப்பலகை என்பதறியலாம். மலேசியா 'நயனம்' இதழாளர் இராஜகுமரன் இணையம் என்ற சொல்லை 1996-ல் தந்தார். வைய விரிவு வலை, உலாவி, பின்னூட்டு, தொடுப்பு, ஒருங்குறி (இராமகி), மட்டுநர் (moderator), வெள்ளுரை (plain text), செழியுரை (rich text), திரட்டி, ... போன்ற கலைச்சொற்கள் கூகுள்குழுக்கள், பதிவுலகில் பரவிவிட்டன.

முகுந்தராசு அளித்த எ-கலப்பை, திஸ்கி, யூனிக்கோடு இரண்டையும் உள்ளடக்கிய உமர்தம்பியின் தேனீ எழுதுரு, ஏதோவொரு குறியில் இருப்பதை ஒருங்குறிக்கு மாற்ற சுரதாவின் பொங்குதமிழ் மாற்றி, ... வலைப்பதிவு எழுதுவதை மக்களுக்கு எடுத்துச் சென்றன. இணையம் உருவாவதன் முன்னமே கோபர் வழங்கியில் சங்க, பிற்கால இலக்கியங்களைத் தந்த கொலோன் பல்கலை மால்ட்டனின் முயற்சி முன்னோடியானது. இணையப் பல்கலை, மதுரை முன்னியம், தமிழ்மரபு அறக்கட்டளை, ஈழ நூல்களை அழிவில் இருந்து காக்கும் நூலகம் திட்டம், எண்மியத் (digital) தேவாரம் (புதுச்சேரி), அகராதிகளுக்கு சிக்காகோ பல்கலை ரோஜா முத்தையா தளம், ... தமிழாய்வில் தினமும் பயன்படும் தளங்கள். இணையப் பல்கலை ஒருங்குறிக்கு மாறவிருப்பதாக அறிகிறோம். 2003-ல் தொடங்கிய விக்கிபீடியா, கலைச்சொல் விக்கினரி உபயோகமான தரவுகளைத் துழாவுவோருக்கு உடனுக்குடனே அள்ளித் தருகின்றன. இவை எல்லாமும் ஒருங்குறியில் இயங்குவதால் தேடிகளில் துழாவல் எளிது, எனவே சிங்கப்பூர் தன் ஒரே குறியேற்பாக அதனை அறிவித்துள்ளது.

3.0 இந்திய மொழிகளின் முதல் வலைத்திரட்டி 'தமிழ்மணம்'

”தமிழ்மணத்தில் என்னையும்வை என்றேனே – அவன் தனிமனத்தில் இருநினைவா என்றானே” என்று காதலி கூற்றாக அன்றே சொல்லிப் போந்தார் புரட்சிக்கவிஞர். கணினி முன்னால் அமர்ந்தால் ஆகும் காலச் செலவைப் பார்த்தால் அந்தத் தீர்க்கதரிசன வாக்கின் உண்மை புலப்படும். 2003 முதல் தமிழ் வலைப்பூக்கள் மலர்ந்தன. பதிவுகளுக்காக நியூக்லியஸ் நிரலைத் தமிழ்ப்படுத்திய காசி ஆறுமுகம் தமிழ்மணம் (<http://tamilmanam.net>) திரட்டியை நிர்மாணித்தார். வலைத்திரட்டிகள் மாற்று ஊடகமாக இணையத்தை நிலைநிறுத்தின. சிற்றிதழ்களுக்கும், தனி வலைப் பதிவுகளுக்கும் தனி வாசகர் வட்டத்தைத் திரட்டிகளே நல்கின. தமிழ்மணம், மறைந்த தேன்கூடு, தமிழ்வெளி திரட்டிகளும், தமிழிஷ் பதிவிணைப்பு தளமும் வலைப்பதியும் புதுமுகங்களுக்கு உற்சாகம் ஊட்டுவன. ஐந்து ஆண்டுகளாய் இயங்கும் தமிழ்மணத்தில் 6000 பதிவுகளும், நாளும் 300 இடுகைகளும் வெளிவருகின்றன, ஒரு நாளுக்குச் சுமார் 2500 மறுமொழிகளையும் திரட்டுகிறது.



தமிழின் - தமிழ்மணத்தின் வலைப்பதிவு வளர்ச்சி (தொடக்க காலம்)

பரந்து கிடக்கும் உலகத் தமிழரை இணைத்துப் பயன்படும் தகவல்களை உடனுக்குடன் பரிமாறத் திரட்டிகள் துணை இன்றுண்டு. பிரச்சினைகளைச் சுயநலம், வணிக நோக்கம் இன்றி எடுத்துச் சொல்லும் செய்தி ஊடகங்கள் தமிழிலில்லை. அனைத்து நாடுகளிலும் தமிழர்கள் சிறுபான்மையர் என்பதால் அரசுடகங்களின் ஆதரவில்லை. PBS தொலைக்காட்சி, NPR வானொலி போன்றவை ஏற்படுத்தும் நிதிவசதியும் இல்லை. இக்குறையை நீக்க வலைப்பதிவுலகம் நிறைய உதவும். ஒரு பிரச்சினைக்கு எல்லாப் பரிமாணங்கள், பல்தரப்பு விவாதங்களை அளிக்கும் பதிவுகளால் உண்மையைப் புரிந்துகொள்ள முடிகிறது. தமிழ் செம்மொழி இலக்கியங்களைக் கற்கும்

மாணவர்களுக்குப் இலக்கண, இலக்கியங்களை விளக்கப் பதிவுலகில் இயலுகிறது. இணையத்தில் பாரிஸ் பேரா. செவ்வியாரும் நானும் 15 வருடமாகத் தொடர்ந்து தமிழ்பற்றி வலையாடுகிறோம்.

4.0 தரவுத்தளங்களில் எழுத்து வரிவடிவங்கள்

4.1 எதிர்காலத்தில் இந்தியாவில் எழுத்துப்பெயர்ப்பும், கணிமொழிபெயர்ப்பும்

ஐரோப்பிய நாடுகள் காலனி ஆட்சியில் சிக்கவில்லை. சீன நகரும்-அச்சு நுட்பம் ஐரோப்பாவில் அறிமுகமாகி குட்டன்பர்க் போன்றோரால் பதிப்பகங்கள் பெருகின. ஒவ்வொரு மொழியும் விவிலியம், சர்ச் போதனை நூல்களாக முதலில் வளர்ந்தது. கல்விப் பரவலாக்கம் மக்களிடையே மறுமலர்ச்சி, விழிப்புணர்ச்சி, விடுதலை உணர்வுகளை உள்வாங்கித் தனிநாடுகளைக் கட்டமைத்தன. அந்த இயல்பான வளர்ச்சி காலனி ஆதிக்கத்தால் இந்தியாவில் தடைப்பட்டது. ஓரளவுக்கு இந்திய நடுவண் அரசு வேற்றுமைகளை அங்கீகரிப்பதன் வடிகால்களாக மொழிவாரி மாகாணங்களை அனுமதித்தது. இந்திய ஒருமைப்பாட்டு நிலையான அரணுக்கு இந்திய ரூபாய் நோட்டீஸ் காணப்படும் எல்லாத் தேசிய மொழிகளுக்கும் சமமான ஒரே அந்தஸ்து நடைமுறையில் வழங்கப்படல் வேண்டும். 19-ஆம் நூற்றாண்டின் 'தாய்மொழி' (mother tongue) கோட்பாட்டின் விளைச்சல் மொழிவாரி மாநிலங்கள் என்ற தகுதியேனும் கிடைத்தது எனலாம். இனி, 'தாய்எழுத்து' (mother script) பயன்படும் வகையைப் பார்ப்போம்.

தமிழ்நாட்டிலிருந்து கல்லாத மக்களும், தொடக்கக்கல்வியே உடையோரும் வடஇந்தியாவில் வாழச் சென்றால் இந்தியைத் தானாகக் கற்றுக்கொள்கின்றனர். அதேபோல், தமிழ்நாட்டுக்கு வேலைக்கு வருவோரும் தமிழைப் பேசுகின்றனர். இவர்கள் செல்பேசிகளில், கணியுலா மையங்களில் அவரவர் 'தாய்எழுத்திலே' மற்ற மாநில எழுத்துக்களைப் படிக்கத் தொழில்நுட்பம் உதவுகிறது. மத்திய, மாநில அரசுகளின் மடலாடல், தகவல், அரசாணைகள் அந்தந்த மாநில எழுத்துப்பெயர்ப்பிலும், மொழிபெயர்ப்பிலும் கணிநுட்பங்களால் தானியங்கியாக உருவாகும் நாள் தொலைவில் இல்லை. நிதி, கணிஞர்களை அளித்து கூகுள் போன்ற நிறுவனங்களுடன் அரசு வழிகாட்டினால் இந்திய மொழிகளுக்கு இடையில் (உ-ம்: தமிழ் < > இந்தி) கணிமொழிபெயர்ப்பு வசதி நல்ல வளர்ச்சிபெறும். நகர்கணிகள், செல்பேசிகளில் தமிழ் இல்லாத பட்சத்தில் தரக் கட்டுப்பாட்டுடன் ஆங்கில எழுத்தில் காட்டும் நிரலிகள்

வேண்டும். உத்தமம் – தமிழ்நாடு அரசு நிபுணர்கள் கூடி, தமிழ் > ஆங்கிலம் எழுத்துப்பெயர்ப்பை (transliteration) நிர்ணயிக்க வேண்டும். உதாரணமாக, உயிரெழுத்து இரண்டின் நடுவே உள்ள ககர எழுத்தைத் தவறாக 'g' என்று எழுத்துப்பெயர்ப்பதை இணையத்தில் காணலாம். மொழியியற்படி இது பொருந்தாது: Intervocalical க் in Tamil has a voiceless fricative sound, and usually it is rendered as k or h in world's languages. அழகு, முருகன், அகம், தூரிகை போன்றனவற்றின் உராய்வொலி azaku, murukan, akam, thuurikai அல்லது azahu, muruhan, aham, thuurihai என்று பெயர்க்கப்படல் தமிழின் ஒலிப்புக்குச் சிறப்பு. அன்றேல் கன்னடம், தெலுங்கு ஒலிப்பாகி, தமிழின் தன்மையை இழக்கும். எழுத்துப்பேர்ப்புத் தர-உறுதிப்பாட்டை இணைய மாநாடுகளில் தமிழ்நாட்டு அரசாங்கம் அறிவிக்கவேண்டும். தமிழ் எழுத்து அறியாத் தமிழ்க் குழந்தைகளும், செம்மொழி கற்க வரும் மேலை நாட்டாரும் தமிழின் சிறப்பு ஒலிகளை அறிந்துகொள்ளக் கணினியில் சரியான ரோமன் எழுத்துப்பெயர்ப்பு துணைநிற்கும்.

4.2 எதிர்காலத்தில் எழுத்துச் சீர்மையும், தனித்தமிழ் எழுத்தில் மீக்குறி ஏற்புமும்

தமிழ் எழுத்தைக் கல்லாத மக்கள் லட்சக் கணக்கில் உள்ளனர். ஆங்கிலக் கல்வியே கற்கும் பல மில்லியன் மாணவர்களும் தமிழ்நாட்டிலும், வெளியேயும் வாழ்கின்றனர். அரசியல்வாதிகள், தமிழாசிரியர், பத்திரிகைத்துறை, ... என்று சிறு விழுக்காட்டினருக்கே தமிழ் வருமானம் தரும் தொழிலாக உள்ளது. ஏனையோர் தமிழை விரும்பிப் படிக்கத் தமிழ் எழுத்தின் வரிவடிவம் சீர்மையுற வேண்டும். எழுத்துச் சீர்மையின் தேவையை விளக்கும் அறிஞர் வா. செ. குழந்தைசாமியின் சொற்பொழிவை தமிழ் இணையப் பல்கலை வைத்துள்ளது: <http://www.tamilvu.org/esvck/index.htm>

இலங்கை, இந்தியத் தமிழரின் அடுத்த தலைமுறை தமிழில் நுண்பதிவிட (microblogging e.g. Twitter) உயிர்மெய்ச் சீர்மை அவசியம். ஒருங்குறியின் சிறப்பம்சத்தை இங்கே குறிப்பிடலாம்: நன்னூல் இலக்கணம் சொல்வதுபோல சார்பெழுத்தாக, பிணை விலங்கை உடைத்துத்தான் உ/ஊ உயிர்மெய்கள் ஒருங்குறியில் ஏற்றப்பட்டுள்ளன. உ/ஊ உயிர்மெய்களைப் பிரித்தும் விரும்புவோர் இணையத்தில் படிக்கலாம், எழுதலாம் என்ற அரசாணை வந்தால் தமிழ் கற்பித்தல் எளிமையாகும். செம்மொழி படிக்கத் தமிழினத்தில் பிறவாதோரும் வருகின்றனர். அவர்களுக்கும் உதவும்.

வீரமாமுனியின் பெரும் சீர்திருத்தமும், 1978-ல் அரசு கொண்டந்த சிறு மாற்றமும் செய்த நன்மை குன்றின் மேலொளிரும் விளக்கு. இனிக் குறைந்தபட்சச் சீர்மையாக உ, ஊ உகரக் குறிகள் திருத்தம் தேவை. பழக்கமான ஊகார கிரந்தக் குறியும், அழகியல் பொருத்தப்பாடுடைய, வரிநீளத்தை அதிகரிக்காத உகரக் குறிக்கான பரிந்துரை (முதல் 4 மெய்யெழுத்து வரிசை, அ-ஓ உயிர்மெய்கள்) எடுத்துக்காட்டு:

ஃ	அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஓ
க்	க	கா	கி	கீ	க்ய	கௌ	கெ	கே	கை	கொ
ங்	ங	ஙா	ஙி	ஙீ	ங்ய	ஙௌ	ஙெ	ஙே	ஙை	ஙொ
ச்	ச	சா	சி	சீ	ச்ய	சௌ	செ	சே	சை	சொ
ஞ்	ஞ	ஞா	ஞி	ஞீ	ஞ்ய	ஞௌ	ஞெ	ஞே	ஞை	ஞொ

மேலதிக விவரமாக, தமிழ் நெடுங்கணக்கும், சீர்மையில் கட்டுரையும்: <http://nganesan.blogspot.com/2009/08/ezhuttu-korppu.html>

ஐரோப்பிய மொழியினர் உலகின் எல்லா எழுத்துக்களையும் ரோமன் எழுத்துக்கள் மீது மீக்குறிகளை (diacritical marks) ஏற்றிக் காட்டுகின்றனர். (அ) அம்முறை தனித்தமிழ் இலக்கணம் கூறும் 18 மெய்யெழுத்துக்களின் மீதும் ஏற்பட்டால் கிரந்த எழுத்துக்கு ஒரு மாற்றுவழியாகும். (ஆ) ஒலிப்பு பிசகாது இருப்பதால் துணைக்குறிகளின் தேவை - பிறமொழி எழுத்துக்களைத் தமிழ் எழுத்தில் பெயர்க்க - இருக்கிறது. இதற்குத் தரக்கட்டுப்பாடும், யூனிக்கோடு வரைபொறி (rendering engine) அனுமதிக்க வைக்கவும் உத்தமம்-இணையப் பல்கலை குழுவமைக்க வேண்டும். துணைக்குறிகளில் பொருத்தமானவற்றைத் தேர்ந்து அறிவிக்கலாம்.

5.0 முடிபுரை

இதுகாறும் தமிழ்க் கணினி, வலைப் பதிவு, தரவுத்தள வளர்ச்சிகளின் முக்கிய மைல்கல்களைப் பார்த்தோம். தமிழ் ஆய்வு சிறக்கவும், கற்பித்தல் எளிமையாவதற்கும் தேவையான (1) எழுத்துப்பெயர்ப்பு விதிமுறைகள் (2) எழுத்துச் சீர்மையில் குறைந்த பட்ச அரசாணை (3) தமிழ் எழுத்துக்களில் ஐரோப்பிய எழுத்துக்கள் போல மீக்குறி ஏற்றம் பற்றிய அறிமுகத்தையும் இக்கட்டுரை வழங்கியுள்ளது. விரிவாக, என் வலைப்பதிவில் (<http://nganesan.blogspot.com>) கருத்துத் தெரிவிக்கலாம்.