

# தமிழ்க் கணிப்பில் மறைந்துள்ள சில சிக்கல்களும் சவால்களும் (Hidden problems and challenges in Tamil Computing)

சு. சீனிவாசன், அ. ஜேம்ஸ்  
(தமிழ் இணையப் பல்கலைக்கழகம், சென்னை)

&

எஸ். அனந்தநாராயணன், கே.ஆர்.எஸ். நாராயணன்  
(இந்திராகாந்தி அணு ஆராய்ச்சி மையம், கல்பாக்கம்)

## முன்னுரை

தமிழும் ஆங்கிலமும் இடையிழைந்த எழுத்துவகையை(font) ரோமன் வரிவடிவுக்குக்(script) கொண்டுபோதல், பின்னர் திருத்தம் மேற்கொண்டு மீண்டும் அதனை மூல எழுத்துவகைக்குக் கொண்டுவருதல், தமிழ்-ஆங்கில எழுத்துருக்கள்(characters) ஒருசேர இடம்பெறும் இருமொழி எழுத்துவகையில்(bilingual font) அச்சிடப்பட்ட பக்கங்களை அடையாளம் காணுவதில்(character recognition) ஏற்படும் சிக்கல்கள் மற்றும் அவற்றைத் தவிர்ப்பதற்குரிய சில வழிகள், உரையிலிருந்து பேச்சு உருவாக்கம்(text-to-speech) செய்யும்போது சொற்களின் தொடக்கத்தில், இடையில், இறுதியில் வருகின்ற குறில்,நெடில் ஒலிகளின் லிப்பளவுகள்(articulation duration), ஈற்றயல்(penultimate) எழுத்துக்களில் ஏற்படும் ஒலிச்செறிவு(intensity) மாற்றங்கள், பேசுவதைச் சொற்களாக மாற்றும் மென்பொருள்(dictation software) உருவாக்கத்திற்குத் தேவையான தமிழ்ச் சொற்களஞ்சியத்தை வடிவமைப்பதில் எதிர்கொள்ள வேண்டிய சிக்கல்கள் ஆகியவற்றை இந்தக் கட்டுரை விளக்க முற்படுகிறது.

## எழுத்துப்பெயர்ப்பில் ஏற்படும் சிக்கல்கள்

இணையத்தளத்தில் ஆங்கிலத்தோடு பிறமொழிகளில் தகவல்களைப் பரிமாறிக்கொள்ள இருமொழி எழுத்துவகைகள் பெரிதும் உதவுகின்றன. தமிழர்களால் பரவலாகப் பயன்படுத்தப்படும் இருமொழி எழுத்துவகைக்குச் சான்றாக TAB-ANNA, TSCII ஆகியவற்றைக் குறிப்பிடலாம். இவற்றைக்கொண்டு தமிழ், ஆங்கில எழுத்துருக்களை ஒருசேரக் காணும் வாய்ப்புக் கிட்டுகிறது. பிற நாட்டார் தேவைக்காகத் தமிழ்நூல்களை ரோமன் முறையில் எழுத்துப்பெயர்ப்பு(transliteration) செய்யவேண்டிய தேவை எழுகிறது. தஞ்சைத் தமிழ்ப் பல்கலைக்கழகத்தில் பின்வரும் எழுத்துப்பெயர்ப்பு முறை பின்பற்றப்படுகிறது.

அ(a), ஆ(aa), இ(i), ஈ(ii), உ(u), ஊ(uu), எ(e), ஏ(ee), ஐ(ai), ஒ(o), ஓ(oo), ஔ(au), ஃ(akf)  
க்(k), ங்(ng), ச்(c), ஞ்(nj), ட்(T), ண்(N), த்(t), ந்(nd), ப்(p), ம்(m),  
ய்(y), ர்(r), ல்(l), வ்(v), ழ்(z), ள்(L), ற்(R), ன்(n)

ரோமன் முறையில் பெயர்க்கப்பட்ட பகுதியில் திருத்தம் மேற்கொண்ட பின்னர் தமிழ் எழுத்துருவுக்குச் செல்லும்போது சிற்சில இடங்களில் சிக்கல்கள் தோன்றுகின்றன. காட்டாக, கீழ்க்காணும் வாக்கியத்தை எடுத்துக்கொள்ளுங்கள். 'அஓஉ' ஆகிய மூன்றும் சுட்டெழுத்துக்கள் ஆகும். இங்கு 'அஓஉ' என்பது ரோமன் எழுத்துப்பெயர்ப்பில் 'aiu' ஆகிறது. அடுத்து மீண்டும் தமிழ் எழுத்துருவுக்கு வரும்போது 'aiu' என்பது 'ஐஉ' என்றாகிறது. இவ்வாறே

அஉ(au) என்பது ஓள எனத் தமிழில் மாற்றம் அடைகிறது. செய்யுட்களில் பொருளை வேறுபடுத்திக்காட்டத் தனி உயிர் சீர் நடுவே வருவதுண்டு. (எ.கா.) என்ஐமுன் (என் தலைவன் முன்னால்: குறள்-771 ). கீழ்க்காணும் சொற்களைத் தமிழ் வடிவிலிருந்து ரோமன் வடிவிற்கும், பின்னர் ரோமன் வடிவிலிருந்து தமிழ் வடிவிற்கும் மாற்றும்போது பிழைகள் ஏற்படுகின்றன.

மனஇருள்	=>	மனைருள்
மனஉளைச்சல்	=>	மனெளைச்சல்
என்அன்பே	=>	எனன்பே
என்உயிரே	=>	எனுயிரே

(இங்கு என்அன்பே, என்உயிரே ஆகியன மூவசைச் சீர்களாகவும் எனன்பே, எனுயிரே ஆகியன ஈரசைச் சீர்களாகவும் கொள்ளப்படும்.)

மேற்குறிப்பிட்ட பிழைகள் நிகழாவண்ணம் தவிர்க்க, சொல் இடையில் வருகின்ற உயிரெழுத்தை ரோமன் வடிவிற்குப் பெயர்க்கும்போது அடிக்கோடு(underscore) ஒன்றை உயிரெழுத்துக்கு முன் சேர்த்துக்கொள்ளுவது பயனளிக்கிறது. (எ.கா)

மனஇருள்	=>	mana_iruL
என்அன்பே	=>	en_anpee
தஇப	=>	ta_ipa

(இங்கு தஇப என்பது தமிழ் இணையப் பல்கலைக்கழகத்தின் எழுத்துச் சுருக்கமாகும்)

அடுத்து, இருமொழி எழுத்துவகையில் ஒள்,ஒளி, தெள்,தெளி ஆகியவற்றிலுள்ள புள்ளி மற்றும் இகரக் கொம்பு ஆகியன தனித்து வருவதால், இவை கவனமாகக் கையாளப்படல் வேண்டும். அன்றேல் அவை பெயர்ப்பு நிரல்களால்(conversion programs) ஒள்காரமாகவும், தெள்காரமாகவும் திரிபடைய வாய்ப்புள்ளது. இவ்வகையான பிழைகள் நிகழாவண்ணம் தவிர்க்க, தொடராக வருகின்ற நான்கு எழுத்துருக்களை நிரலில் அலசி ஆராய்தல் அவசியமாகிறது.

இருமொழி எழுத்துவகையில் தமிழோடு ஆங்கிலம் இடையிழைந்து வருகின்ற வாய்ப்பு உள்ளது. (எ.கா) உரைநடையில் எழுத்து(letter), சொல்(word) ஆகியவற்றைக் கவனமாகக் கையாள வேண்டியிருக்கிறது. மேற்குறிப்பிட்ட சொற்றொடரை ரோமன் வடிவிற்குக் கொண்டுபோய் பின்னர் தமிழ் வடிவிற்கு வரும்போது மூலத் தொடரிலுள்ள letter, word ஆகிய சொற்கள் முறையே லெத்தெர், wஒர்ட் எனத் திரிபடைகிறது. இத்தகைய பிழை நேரிடாவண்ணம் தவிர்க்க மூலத்திலுள்ள ஆங்கிலச் சொற்களை ரோமன் வடிவிற்குப் பெயர்க்கும்போது அவற்றை ஏதாவது ஒரு சிறப்பு அடைப்புக் குறிக்குள்(tag) இருவது நலம் பயக்கும். இதற்காகக் கோண அடைப்புகளைக்(< >) கூட ஒருவர் பயன்படுத்தலாம். பின்னர் பெயர்ப்பு நிரல்களில் இந்த அடைப்புக் குறிகளை அடையாளம் கண்டு ஆங்கிலச் சொற்கள் திரிபடையாத வகையில் காத்துக் கொள்ளலாம்.

எழுத்துருக்களை அடையாளம் காணுவதில் ஏற்படும் சிக்கல்கள்

தமிழும் ஆங்கிலமும் சேர்ந்து வருகின்ற பகுதிகளில் எழுத்துருக்கள் சிலவற்றை அடையாளம் காணுவதில் சிக்கல் எழுகின்றது. இந்தப் பகுதிகள் கையால் எழுதப்பட்டனவாகவோ அச்சிடப்பட்டனவாகவோ இருக்கலாம். சான்றாக, ரோமன் வடிவில் இடம்பெறும் o, I(small-L), s(small-S), z ஆகிய எழுத்துக்களை 0(number-0), I(capital-i), 5(number-5), 2(number-2)

ஆகியவற்றிலிருந்து வேறுபடுத்தி அறிவதில் சிக்கல் எழுகிறது. இதேபோல கீழ்க்காணும் தமிழ் எழுத்துருக்களை வேறுபடுத்திக்காட்டக் கூடுதல் முயற்சி தேவைப்படுகிறது.

௨-எ; ௩-ன; ௪-ச; ௫-ல; ௬-ஐ; ௭-ஓ; ௮-ஈ; ௯-மு;

தமிழிலும் ஆங்கிலத்திலும் இடம்பெறுகின்ற எழுத்துருக்களில் சில தோற்றத்தில் தமக்குள் ஒத்திருக்கின்றன. அவையாவன:

L (upper case)	-	ட (தமிழ் Ta)
w (lower case)	-	ய (தமிழ் ya)
U (upper case)	-	ப (தமிழ் pa)
y (lower case)	-	பு (தமிழ் pu)

இவற்றை அச்ச ஊடகத்தில் பிரித்துக்காட்டுவது அவசியமாகின்றது. தமிழிலுள்ள டகரம் சற்றுக் கிடையாக நீள்வதாகவும், ஆங்கிலத்திலுள்ள L சற்று நெடுக்காக உயர்வதாகவும் காட்டவேண்டிய தேவை எழுகின்றது. இத்தகைய மாற்றங்களை இருமொழி எழுத்துவகையில் பகுத்தினாலன்றி அவற்றைப் பிரித்தறிவது கடினமாகும்.

உரையிலிருந்து பேச்சு உருவாக்கம் செய்வதில் கவனிக்கத் தக்கவை

பல்ஊடக(multimedia) வசதியோடு WaveStudio பயன்பாட்டு நிரல்கொண்டு தமிழ்ச் சொற்களில் இடம்பெறுகின்ற எழுத்தொலிகள் பகுப்பாய்வு செய்யப்பட்டன. அகர ஒலிப்பு நேரம், மகர ஒலிப்பிற்குரிய நேரத்தில் மூன்றில் ஒரு பங்குதான் இருக்கின்றது. இத்தகவல் குறில் எழுத்துக்களின் ஒலிப்புநேரம் எல்லா எழுத்துக்களுக்கும் ஒரேயளவு இருப்பதில்லை என்பதைக் காட்டுகிறது. குறில், நெடில் எழுத்துக்களுக்கான ஒலிப்புநேரம் சொல்லின் தொடக்கம், இடை, இறுதி நிலைகளைப் பொருத்தும் மாறுகிறது.(காண்க அட்டவணை-1).

அட்டவணை 1. குறில், நெடில் ஒலிப்பளவுகள்

பண்பு	ஒலிப்பு நிலை		
	சொல் தொடக்கம்	சொல் இடை	சொல் இறுதி
1. குறில் ஒலிப்பளவு (சராசரி)	145 (மி.செ)	163 (மி.செ.)	281 (மி.செ.)
2. நெடில் ஒலிப்பளவு (சராசரி)	230 (மி.செ)	236 (மி.செ.)	327 (மி.செ.)
3. மெய் ஒலிப்பளவு (சராசரி)	-	161 (மி.செ.)	195 (மி.செ.)
4. ஆ/அ சார்பு ஒலிப்பு விகிதம்	1.74	1.69	1.14
5. ஈ/இ சார்பு ஒலிப்பு விகிதம்	1.90	1.57	1.39
6. ஊ/உ சார்பு ஒலிப்பு விகிதம்	1.98	1.45	1.33
7. ஏ/எ சார்பு ஒலிப்பு விகிதம்	1.55	1.45	-
8. ஓ/ஔ சார்பு ஒலிப்பு விகிதம்	1.53	1.56	-

மேலும் சொற்களின் ஒலிப்பளவு சில நியதிகட்கு உட்பட்டு இயங்குவதையும் அறிய முடிகிறது. அவற்றுள் சில கீழே தரப்பட்டுள்ளன.

1. சொல் நீளம் கூடும்போது, தொடக்க எழுத்தின் ஒலிப்பளவு குறைகிறது. (எ.கா) (த)லை, (த)லைவி, (த)லைவிதி; (பா)தி, (பா)திரி, (பா)திரியார்;

இது பேருந்தில் இருவர் அமரும் இருக்கையை மூவரோ அல்லது மூவர் அமரும் இருக்கையை நால்வரோ பகிர்ந்துகொள்ளும் நிலைக்கு ஒப்பானதாகும்.

2. சொல் நீளம் கூடும்போது, தொடக்கத்தில் இடம்பெறும் குறில் இணையின்(pair) ஒலிப்பளவு குறைகிறது. (எ.கா) (எழு)து, (எழு)பது

3. பின்வரும் இணைகளில், முதற் சொல்லின் தொடக்கத்தில் இடம்பெறும் உயிர்க்குறிலின் ஒலிப்பளவு கூடுகிறது.  
(அ)றி, (அ)றிவு ; (மு)டி, (மு)டிவு ;  
(ச)வம், (ச)வரம் ; (அ)கம், (அ)கரம் ;

சொற்களிலுள்ள எழுத்தொலிகளின் ஒலிப்பளவு மாறுபடுவதைப்போல அவற்றின் ஒலிச்செறிவும் மாறுதல் அடைகின்றது. காட்டாக வனம், நவம் ஆகிய சொற்களை ஒலிக்கும்போது ஈற்றெழுத்துக்கு முன்னுள்ள எழுத்தொலி வலிமை பெறுகிறது. இந்த வேறுபாடுகளை அறிந்து ஒலியின் வலிமையைப் பேச்சு உருவாக்கத்தின்போது கூட்டவும் குறைக்கவும் முயற்சி மேற்கொள்வது அவசியமாகின்றது.

மேலே கோடிட்டுக் காட்டியுள்ள ஒலிப்பளவு, ஒலிச்செறிவு ஆகிய பண்புகளை மென்பொருளில் எதிரொளிக்கச் செய்வது ஒரு சவாலான பணியாகத் தோன்றுகிறது.

பேச்சிலிருந்து சொல் உருவாக்கம் செய்வதில் எதிர்கொள்ள வேண்டிய சவால்கள்

பேசுவதைச் சொற்களாக மாற்றும் மென்பொருள்(dictation software) உருவாக்கத்திற்குத் தேவையான தமிழ்ச் சொற்களஞ்சியத்தை வடிவமைப்பது சிக்கல் நிறைந்த பணியாகத் தோன்றுகிறது. ஆங்கிலத்தில் மொழிவதைத் திரையில் எழுதிக்காட்டவோ ஆவணத்தில்(document) சேர்க்க உதவவோ Dragon Naturally Speaking(DNS) முதலிய பயன்பாட்டு நிரல்கள் உதவுகின்றன. ஒலிவாங்கியின் (microphone) முன்னர் பேசும்போது, நாம் எழுப்பும் ஒலிகள் சொற்களாகவோ சொற்றொடர்களாகவோ பதிவாவதில்லை. மாறாக, அவை தொடரும் ஒலி அலைகளாகவே(continuous stream of sounds) பதிவாகின்றன. இவற்றிலிருந்து பொதுவான ஒலிக்கூறிப்புகளையும்(sound patterns), ஒலியன்களையும்(phonemes) அடையாளம் காணும் பணியை DNS போன்ற நிரல்கள் செய்கின்றன. இத்தகைய ஒலிக்கூறிப்புகளைச் சொற்களாக மாற்றுவதற்கு ஒலித் தரவு(acoustic data) மற்றும் மொழித் தரவு(language data) அடிப்படையாக அமைகின்றன. ஒருவரின் உச்சரிப்பை(pronunciation) அறிவதிலிருந்து கருவியின் செயல்பாட்டிற்குத் தேவையான ஒலியன்கள் பற்றிய தகவல், சொற்களுக்கு இடையே விடப்படுகின்ற நேரம், ஒலிஅதிர்வு(pitch) முதலிய தரவுகள் கிடைக்கின்றன. இதற்குக் கருவிமுன் உட்கார்ந்து சில சொற்களையும் தொடர்களையும் உச்சரித்துப் பயிலுவது அவசியமாகிறது. இந்தப் பயிற்சியின்(training) மூலம் ஒருவர் எழுப்பும் ஒலி உருவகைகளை(patterns) சொற்களஞ்சியத்தில் பதிவு செய்யப்பட்டிருக்கும் ஆயிரக்கணக்கான சொற்களோடு பொருத்திப் பார்க்கும் திறனைக் கருவி பெறுகிறது. தமிழில் பேசுவதை அடையாளம் காண்பதற்கு 40 ஒலியன்கள் போதுமானவை என மொழியியல் அறிஞர்கள் உரைக்கின்றனர்.

பயிற்சி மேற்கொண்ட பிறகும் ஒலித் தகவல்களைச் சரிவரப் புரிந்துகொள்ள இயலாத நிலையில், மொழித் தரவுகளைச் சார்ந்து கருவி இயங்க வேண்டிய சூழல் உருவாகிறது. இதற்குச் சொற்களஞ்சியங்கள்(அகராதிகள்) உதவுகின்றன. பெயர்(noun), வினை(verb) ஆகியவற்றை அடிப்படையாகக்கொண்டு உருவாக்கப்படும் அனைத்துச் சொற்களையும்(derived

words) அகராதியில் இடம்பெறச்செய்வது இயலாத செயலாகும். சில சொற்களை இயங்கு நேரத்தில்(run time) உருவாக்க வேண்டிய தேவையும் எழுகிறது. காட்டாக, ஆங்கிலச் சொற்களில் பன்மை விசுவயைச் சுட்ட s என்ற எழுத்து சொல் இறுதியில் சேர்க்கப்படுகிறது. (எ.கா.) boys, girls. இதுபோன்றே ஆறாம் வேற்றுமை உருபைக்(possessive case) காட்ட மேற்கோள் குறியோடு(apostrophe) s என்ற எழுத்தும் சேர்க்கப்படுகிறது. (எ.கா.) Newton's rings; ஆங்கிலச் சொற்களில் படர்க்கை, ஒருமையைக் காட்டும் நிகழ்கால வினைமுற்றிலும் s என்ற எழுத்து இறுதியில் சேர்க்கப்படுகிறது. (எ.கா) He plays; it works;

இறந்தகால வினைமுற்றைச் சுட்ட ed எனும் இரண்டு எழுத்துக்கள் வினையடியோடு சேர்க்கப்படுகின்றன. (எ.கா.) walked. slowed; இத்தகைய விதிமுறைகளுக்குக் கீழ்வராத திரிபு வடிவங்களே அகராதியில் காட்டப்படுகின்றன. (எ.கா.) went, slept. இது போன்றே men, teeth முதலிய பன்மையைச் சுட்டும் சொற்கள் தனித்த பதிவாக(separate entry) அகராதியில் இடம்பெறுகின்றன. திணை, பால், எண், இடம், காலம், உருபு ஆகியவற்றால் மாறுதல் அடையும் he, she, it, they, her, hers, his, him, their theirs; be, is, am, are, was, were, shall, should, will, would போன்ற சொற்களையும் அகராதியில் காணலாம்.

தமிழ்மொழியிலும் இத்தகைய ஒழுங்கமை(regular), ஒழுங்கிலா(irregular) வடிவங்களைப் பெயர்ச்சொல், வினைச்சொல் அடிப்படையில் வகைப்படுத்த முடிகிறது. பெரும்பாலான தமிழ்ச்சொற்களில் பன்மையைக் காட்ட 'கள்' விசுவயைச் சேர்க்கப்படுகிறது.

இத்தகைய சொற்கள் சில நியதிகட்கு உட்படுவதையும் காணலாம். மகர ஒற்றுடன் முடியும் சொற்கள் நகர ஒற்றாகத் திரிந்து பன்மை வடிவம் கொள்கின்றன. (எ.கா)

இனம்+கள் = இனங்கள்  
சட்டம்+கள் = சட்டங்கள்  
மரம்+கள் = மரங்கள்

உயர்வு நிமித்தம் பெயர்ச்சொல், வினைச்சொல் ஆகியவற்றோடு 'கள்' விசுவயைச் சேர்ப்பதும் இந்நாளில் வழக்கில் காணப்படுகிறது. (எ.கா.) அவர்கள் வந்தார்கள்.

பெயர்ச்சொற்களோடு வேற்றுமை உருபுகளைச்(case markers) சேர்க்கும்போது சொல் வடிவம் திரிகிறது. இவற்றையும் சில விதிமுறைகளுக்குள் அடக்கிவிட முடிகிறது. காட்டாக இரண்டாம் வேற்றுமை உருபாகிய 'ஐ'யின்(objective case) புணர்ச்சி விதிகளைக் கீழ்க்கண்டவாறு வகைப்படுத்தலாம்.

1. தனிக் குறிலை அடுத்த மெய், உயிருடன் (இங்கு ஐ) புணரும்போது இரட்டிக்கும்.  
எண்+ஐ = எண்ணை; நெய்+ஐ = நெய்யை; கல்+ஐ = கல்லை; எள்+ஐ = எள்ளை;
2. தனி நெடிலை அடுத்த மெய், உயிருடன் (இங்கு ஐ) புணரும் போது இரட்டிக்காது.  
ஆண்+ஐ = ஆணை ; நோய்+ஐ = நோயை; கால்+ஐ = காலை; வாள்+ஐ = வாளை;
3. இ, ஈ, ஏ, ஐ ஆகிய உயிர்களின் முன், உயிர்வந்தால் (இங்கு ஐ) இடையில் 'ய்' தோன்றும்  
கிளி+ஐ = கிளியை; தீ+ஐ = தீயை; மலை+ஐ = மலையை;
4. பிற உயிர்களின் முன் (இ, ஈ, ஏ, ஐ, குற்றுகரம் நீங்கலாக), உயிர்வந்தால் (இங்கு ஐ) இடையில் 'வ்' தோன்றும்  
பலா+ஐ = பலாவை; பூ+ஐ = பூவை; திரு+ஐ = திருவை;

5. குற்றுகரத்தை அடுத்து ஐ வரும்போது, ஈற்று உகரம் கெடும்.  
 விறகு+ஐ = விறகை; காசு+ஐ = காசை; காற்று+ஐ = காற்றை; கதவு+ஐ = கதவை;
6. மகர ஒற்றுடன் முடியும் மூன்று அல்லது மூன்றுக்கு மேற்பட்ட எழுத்துக்களை உடைய சொற்கள் ஐயுடன் புணரும்போது இடையில் 'அத்து' என்னும் சாரியைத் தோன்றும்.  
 இடம்+ஐ = இடத்தை; இன்பம்+ஐ = இன்பத்தை;  
 காலம்+ஐ = காலத்தை; செல்வம்+ஐ = செல்வத்தை;  
 வானம்+ஐ = வானத்தை; பயணம்+ஐ = பயணத்தை;

தமிழில் வினைமுற்றாகவும் எச்சமாகவும் வருகின்ற சொற்கள் திரிபடையும் நிலையை 12 வகைகளாக டாக்டர்.கிரெளல்(Graul) பிரித்திருக்கிறார். (நோக்குக: தமிழ்ப் பேரகராதி, தொகுதி-1) அவற்றின் வினைப்பகுதியாவன :

1. செய்- 2. ஆள்- 3. கொல்- 4. அறி- 5. ஆக்கு- 6. நடு-
7. உண்- 8. தின்- 9. கேள்- 10. கல்- 11. தீர்- 12. நட-

காட்டாக ஆக்கு(வகை-5) என்ற வினைப்பகுதியிலிருந்து ஆக்கினேன், ஆக்குகிறேன், ஆக்குவேன் ஆகிய வினைமுற்றுகள் தோன்றுகின்றன. 'ஆக்கு' என்ற வினைப்பகுதியின் அடிப்படையில் கீழ்க்காணும் வினையடிகள் அமைவதைக் காணலாம்.

அனுப்பு, ஆடு, உறங்கு, எழுது, ஓடு, கீறு, தங்கு, தாண்டு, தூக்கு, தூங்கு, தோண்டு, நிரப்பு, நீந்து, நெருங்கு, பூசு, மடக்கு, மாற்று, வாங்கு, வாடு.

இது போன்று, செய்(வகை-1) எனும் வினைப்பகுதியிலிருந்து எச்சம், முற்று, தொழிற்பெயர், வினையாலணையும் பெயர் எனத் திரிபடையும் சொற்கள் (இணையத் தளத்திலிருந்து இறக்கப்பட்ட ஓராண்டு 'தினபூமி' நாளேட்டுச் செய்திகளிலிருந்து பெறப்பட்டவை) கீழே தொகுத்துத் தரப்பட்டுள்ளன.

செய் செய்க செய்கிற செய்கிறது செய்கிறார் செய்கிறார்கள்  
 செய்கிறேன் செய்கிறோம் செய்கின்றன செய்கின்றனர் செய்த செய்தற்கு  
 செய்ததாக செய்ததாகவும் செய்ததாகவேசெய்ததால் செய்ததால்தான் செய்ததில்  
 செய்தது செய்ததை செய்ததோடு செய்தபடி செய்தபின் செய்தபோது  
 செய்தல் செய்தவர் செய்தவர்கள் செய்தவர்களுக்கு செய்தவர்களை  
 செய்தவருக்கு  
 செய்தவரை செய்தவன் செய்தன செய்தனர் செய்தார் செய்தார்கள்  
 செய்தாரா செய்தால் செய்தாலும் செய்தாள் செய்தான் செய்திருக்க  
 செய்திருக்கிறார் செய்திருந்தனர் செய்திருந்தார் செய்திருந்தாசெய்திருப்பதாக  
 செய்து  
 செய்துகொண்ட செய்துகொண்டவர்கள் செய்துகொண்டனர்  
 செய்துகொண்டார்  
 செய்துகொண்டு செய்துதர செய்தும் செய்துவருகிறது செய்துவரும்  
 செய்துவிட்டதாக  
 செய்துவிட்டது செய்துவிட்டனர் செய்துவிட்டார் செய்துவிட்டு  
 செய்துவிட்டேன் செய்துவிடுவார்கள்  
 செய்துவிடுவீர்களா செய்துவிடுவேன் செய்துவைக்க செய்துவைத்தார்  
 செய்துவைப்பேன்

செய்துள்ள	செய்துள்ளதாக	செய்துள்ளது	செய்துள்ளதை	செய்துள்ளன
செய்துள்ளனர்				
செய்துள்ளார்	செய்துள்ளேன்	செய்துள்ளோம்	செய்தே	செய்தேன்
செய்தோம்	செய்தோருக்கு	செய்பவர்கள்	செய்பவர்களின்	செய்பவர்களுக்கு
செய்பவர்களை				
செய்பவரை	செய்வதற்காக	செய்வதற்கான	செய்வதற்கு	செய்வதற்கும்
செய்வதாக	செய்வதால்	செய்வதில்	செய்வது	செய்வதே
செய்வதை	செய்வார்	செய்வார்கள்	செய்வாரா	செய்வீர்
செய்வீர்களா	செய்வேன்	செய்வோம்	செய்வோர்	செய்வோருக்கு
செய்வோரை				

ஆங்கிலத்தில் இடம்பெறும் go-went-gone; steal-stole-stolen போன்ற ஒழுங்கிலா வினை வடிவங்கள் தமிழிலும் இடம்பெறுகின்றன. இவற்றுக்கு வா, தா, சா, காண் ஆகிய வினைப்பகுதிகளைக் காட்டாகக் குறிப்பிடலாம். இவற்றுக்கான திரிபு நிலைகளை அகராதியில் காட்டுவது அவசியமாகின்றது. (எ.கா) வா-வந்தேன்-வருகிறேன்-வருவேன்

எழுத்தில்(spelling) வேறுபடுகின்ற, ஆனால் ஒலிப்பில் ஒத்த அல்லது சிறிதே மாறுபடுகின்ற சொற்களை அடையாளம் காண்பதிலும் சிக்கல் எழுகின்றது.

(எ.கா.) ஆங்கிலம் : advise-advice; fare-fair; peace-piece-peas; right-rite;  
தமிழ் : அரை-அறை; கனை-கணை; முந்நாள்-முன்னாள்; வலி-வளி;

ஆள் பெயர், ஊர்ப் பெயர் மற்றும் அகராதியில் இடம்பெறாத சொற்களை உள்ளிடுவதற்குச் சொல் சேர்ப்பி(vocabulary builder) போன்ற வசதிகள் DNS முதலிய பயன்பாட்டு நிரல்களில் கொடுக்கப்பட்டுள்ளன. இவற்றைக்கொண்டு எழுத்துக்களைத் தனித்தனியாக உச்சரித்து, அகராதியில் புதிய சொற்களைச் சேர்க்க முடிகிறது. இதற்குப் 'பன்னாட்டுத் தொடர்பு நெடுங்கணக்கு'(International Communication Alphabet) உதவுகின்றது. தந்திச் செய்திகளைத் தொலைபேசி மூலம் அனுப்புவதற்கு இந்த நெடுங்கணக்கு பயன்படுகிறது. காட்டாக A,B,C, முதலிய எழுத்துக்களை அனுப்புவதற்கு Alpha, Bravo, Charlie,.. என உலக வல்லுநர்களால் ஏற்றுக்கொள்ளப்பட்ட சொல் பட்டியல் உருவாக்கப்பட்டுள்ளது. இதனை ஒத்த சொற் பட்டியல் ஒன்று (ஆத்திசூடியின் தொடக்கத்தில் வருகின்ற சொற்கள் போன்று) தமிழுக்கும் உருவாக்கப்படுவது அவசியமாகின்றது.

இதுகாறும் கூறப்பட்டுள்ள சிக்கல்களில் பல தமிழ்மொழிக்கே உரித்தானவையாகும். இவற்றை ஆழ்ந்தும் அகன்றும் ஆராய்வது அவசியமானதாகத் தோன்றுகிறது. சிக்கல் நிறைந்த இப்பணிக்குத் தீர்வு காண்பது காலத்தின் கட்டாயமாகும்.

## முடிவுரை

தமிழும் ஆங்கிலமும் ஒருசேர வரும்பகுதிகளை எழுத்துப்பெயர்ப்புச் செய்யும்போது கவனத்தில் கொள்ளத்தக்கவை, ஒன்றுபோல தோற்றமளிக்கும் எழுத்துருக்களை வேறுபடுத்திக் காண்பதில் ஏற்படும் சிக்கல்கள், உரையிலிருந்து பேச்சுருவாக்கம் செய்யும்போது ஒலிப்பளவு, ஒலிச்செறிவு ஆகியவற்றைச் சொல்லில் இடம்பெறும் எழுத்து நிலையோடு பொருத்திப் பார்த்தல், பேச்சிலிருந்து சொற்களை அடையாளம் காண்பதில் மொழிக்கூறு வகிக்கும் பங்கு ஆகியவை இக்கட்டுரையில் விளக்கப்பட்டிருக்கின்றன.