

యూనికోడ్ ప్రమాణీకరణ - తెలుగుప్రచురణలపై దాని ప్రభావం

- డా. తిరుమల కృష్ణదేశికాచార్యులు, విన్సిపెగ్, కెనడా

ఇది సాంకేతికవిషయమైనా నేను దానిని తెలుగులోనే చెప్పడలచుకొన్నాను. దీనికి కారణం నేను చెప్పబోయే విషయాలకు ఇంగ్లీషుకంటే తెలుగులో చెప్పడమే సహజంగా ఉంటుంది గనుక. ముందుగా సాంకేతికపదాల నిర్వచనాలు:

లిపి - Font

వర్ణము - Character

అక్షరము - Syllable

వర్ణముద్ర లేక అక్షరముద్ర - Grapheme

కంప్యూటర్ తెర - Computer Screen

పత్రము - Document

పరికరణము - Utility

వర్ణప్రమాణపట్టిక - Standard character code table (e., ASCII or Unicode or ISCII code table)

కీఫలకము - Keyboard

భారతీయభాషలలోని వర్ణాల ప్రమాణీకరణకు ఇప్పుడు రెండు ప్రమాణీకరణపద్ధతులు అమలులో ఉన్నాయి. ఒకటి ఇస్కీ (ISCII) పద్ధతి, రెండవది యూనికోడ్ (Unicode) పద్ధతి. ఇస్కీ-పద్ధతిలో భారతదేశంలోని ప్రధానభాషలలోని వర్ణాల నన్నిటిని సమీకరించి, అన్నిభాషల లిపులకు సరిపడునట్లుగా 255 వర్ణాలు గల లిపిలో 128 నుండి 255 వరకుగల స్థానాలు వీటికి కేటాయించడం జరిగింది. ఈ పద్ధతిలో మొదటి 0 నుండి 127 వరకుగల స్థానాలు ఇంగ్లీషుకు, ఇంగ్లీషు అంకెలకు, తదితరవిరామచిహ్నాలకు కేటాయించబడినాయి. ఈ విధంగా చేయడంవల్ల, ముందుగా వాడకందారు ఏ భాషను వాడుతున్నాడో గుర్తించి, ఆ భాషకు తగిన లిపిలోని అక్షరాలను కంప్యూటరు తెరపై చూపేడుతుంది. ఈ అక్షరాలు భాషను బట్టి వేరువేరుగా కనిపించినా, కంప్యూటరుమాత్రం అన్నిభాషలకూ ఒకే ప్రామాణికసంకేతసంఖ్యను పైలులో దాచుకొంటుంది. ఇట్టి పైలును కంప్యూటరు తెరపై చూపేట్టేటప్పుడు, ప్రింటరుపై ప్రింటు చేసేటప్పుడు, పైలులోని ఏయే భాగాలు ఏయే భాషలలో చూపెట్టాలో గుర్తించి, ఆయా భాషలకు తగిన లిపులలో ఆయా భాగాలను ప్రదర్శించడం జరుగుతుంది.

యూనికోడ్ ఇస్కీ-పద్ధతిలో సమీకరించిన వర్ణాలనే ఆయా భాషలకు ప్రత్యేకంగా కేటాయించిన 128 స్థానాలలో పొందుపరచింది. అంటే యూనికోడ్లో ప్రామాణికవర్ణసంఖ్యలు వేర్వేరుభాషలకు వేరుగా ఉంటాయి. ఇండియాలో తయారు చేసే కొన్ని ప్రోగ్రాములు ఇస్కీ-ప్రమాణాన్ని వాడుతున్నాయి. కాని బృహత్తర అంతర్జాతీయసంస్థలైన మైక్రోసాఫ్టు, ఓరకల్ వంటి సంస్థలు యూనికోడ్ నాధారంగా చేసికొని భారతీయభాషలను కంప్యూటరులో గుర్తిస్తున్నాయి. ప్రస్తుతం ఇస్కీ, యూనికోడ్ రెండూ ప్రచారంలో ఉన్నా, బృహత్తరమైన అంతర్జాతీయసంస్థలు సమర్థిస్తున్న యూనికోడ్ ప్రమాణమే చిరకాలం నిలద్రొక్కుకొంటుందనడం సత్యదూరం కాదని నేననుకొంటున్నాను. ఈ మధ్యకాలంలో ఇస్కీలో వ్రాసిన పత్రాలను యూనికోడ్లోనికి, యూనికోడ్లో వ్రాసిన పత్రాలను ఇస్కీలోనికి పరివర్తన చేసికొనే పరికరణాలు కంప్యూటర్ సంస్థలు రూపొందిస్తే బాగుంటుంది. ముఖ్యంగా మైక్రోసాఫ్టు వర్ధవంటి ప్రోగ్రాముల్లో ఇట్లాంటి సౌకర్యాలు కల్పిస్తే అందరికీ ఉపయోగకరంగా ఉంటుంది.

మైక్రోసాఫ్టు విండోస్ 2000లో తొట్టతొలిగా యూనికోడుతో ప్రమాణీకృతములైన భారతీయభాషలను వియోగించుకునే అవకాశం కల్పించింది. కాని విండోస్ XP Professional, వర్డ్ XP వచ్చేవఱకు ఈ అవకాశం తెలుగుదాక రాలేదు. ఇప్పుడు విండోస్ XP Professionalతో మైక్రోసాఫ్టు గౌతమి అనే తెలుగులిపిని పంపిణీ చేస్తున్నది. కాని ఈలిపి అందంగా లేనందు వల్ల తెలుగులో పుస్తకాదిప్రమరణలు దీనితో సాధ్యం కావు. కాని విండోస్ XP Professional మార్కెట్టులోకి రాకముందే మైక్రోసాఫ్టు సహకారంతో విండోస్ 2000 కంటె ముందుగా నున్న విండోసులో పనిచేసే నా పోతనలిపిని పోతన 2000 అనే పేరుతో నేను యూనికోడుతో ప్రమాణీకరించినాను. ఈవిధంగా పోతన 2000 తెలుగుభాషకు రూపొందించిన తొట్టతొలి యూనికోడ్ తెలుగులిపి యైనది. యూనికోడ్ భారతీయభాషలను విండోసులో వాడుకోవడానికిగాను మైక్రోసాఫ్టు ఓపెన్ టైపు లిపిప్రమాణాన్ని ప్రవేశపెట్టింది. ముందుగా విండోసులో ప్రవేశపెట్టిన ట్రూటైపులిపియొక్క విస్తృతరూపమే ఈ ఔపెన్ టైపు. ట్రూటైపులో వర్ణప్రమాణపట్టికలో నున్న వర్ణములను మాత్రమే లిపిలో చేర్చుటకు వీలుగా నుండేది. కాని భారతీయభాషల లిఖితరూపాలు వర్ణమాలములుగాక అక్షరమాలములగుటచే, ఇట్టి ట్రూటైపులిపియందు కేవలము వర్ణములకే స్థానముండు టచే, ట్రూటైపులిపితో తెలుగువంటి భారతీయభాషలను ముద్రించడం చాలా కష్టసాధ్యమైంది. అందుచే వర్ణప్రమాణపట్టిక నియమావళి నుల్లఘించి, ఈ భాషలకు కావలసిన అక్షరాలకూర్చున కవసరమైన అక్షరముద్రలను కూర్చిన ట్రూటైపులిపులు అమలులోనికి వచ్చినవి. కాని ఇవి ప్రామాణికములు కాక పోవుటచే వీనితో కూర్చిన పత్రములను అన్ని ప్రోగ్రాములలో చదువుకొనడానికిగాని, ఇమెయిల్ ద్వారా ఇట్టి లిపులలో వ్రాసిన సమాచారాన్ని పంపడం కాని అసాధ్యమైంది. ఈలోపాలను యూనికోడుతో ప్రమాణీకరించిన పోతన 2000 వంటి ఓపెన్ టైపు లిపులు తీరుస్తున్నాయి. ప్రామాణికవర్ణపట్టికలో నిర్ధరించిన వర్ణములతోబాటు భాషయందలి అక్షరప్రదర్శనకు, ముద్రణకు అవసరమైన అక్షరముద్రలను కూడ పొందుపరచుటకు, వానిని సందర్భానుసారంగా ప్రదర్శించేందుకుగాను అదనపు పట్టికలు నిర్మించే అవకాశం ఓపెన్ టైపులిపిలో కల్పించడంవల్ల పోతన 2000 వంటి లిపులు యూనికోడువర్ణప్రమాణాన్ని పాటిస్తూ తెలుగులిపిని ముద్రించుటకు, ప్రదర్శించుటకు సమర్థవంతము లౌతున్నవి. లేఖకుడు కీఫలకంపై నొక్కిన వర్ణాలను బాషాసంప్రదాయానుసరించి తగువిధంగా మార్పిడి చేసి, కావలసిన ఓపెన్ టైపు లిపినుండి అక్షరపరంపరలను గ్రహించి భాషను తగురీతిలో ప్రదర్శించేందుకుగాను మైక్రోసాఫ్టు యూనిస్కైబు అనే పరికరణాన్ని విండోస్ 2000, విండోస్ XPలతో జతగూర్చింది. ఈ యూనిస్కైబు ప్రోగ్రాము, పోతన 2000వంటి లిపి, రెండింటి సహకారంతో లేఖకుడు కంప్యూటరుపై తెలుగుభాషను యూనికోడుప్రామాణికపద్ధతిలో వ్రాయగల్గుచున్నాడు. విండోస్ 95, 98, MEలతో కూడ యూనిస్కైబును వాడవచ్చునుగాని, అప్పుడి విండోస్ పరిధిలో యూనికోడులో ఇదివరకే వ్రాసిన పత్రాలు చదువడానికి వీలౌతుందిగాని, క్రొత్తగా వ్రాయడానికి వీలుగాదు. ఏమైనను యూనికోడ్ ప్రమాణానికి కట్టుబడిన లిపులలో కంప్యూటరుపై పత్రాలు తయారుచేసుకుంటే, అవి కంప్యూటర్లు మారినా, ప్రోగ్రాములు మారినా ఈ మార్పులకు అతీతంగా తరతరాలకు చదువగలిగేటట్లు ఉంటాయి గనుక, ఇకముందైనా యూనికోడులిపుల నుపయోగించి పత్రాలు వ్రాసి కొనడం ఉత్తమపక్షమని నేను భావిస్తున్నాను.

కాని ఇస్కీ ప్రమాణంగాని, దానిని ఆధారంగా చేసికొని నిర్మించిన యూనికోడు ప్రమాణంగాని అసమగ్రంగాను, లోప భూయిష్టంగాను ఉన్నది. అందుచేత తెలుగులో సంస్కృతభాషలోనున్న పత్రాలను, వేదాలను, నిఘంటువులను కూర్చు చేసి కొనడానికి ఎన్నో ఇబ్బందు లెదురౌతున్నాయి. ఈ క్రింద నా దృష్టికి వచ్చిన లోపాల నన్నిటిని పేర్కొంటాను.

1. అవగ్రహలోపం

సంస్కృతశబ్దాలలో సంధిపరంగా అకారము లోపించినప్పుడు, ఆ అకారముయొక్క ఉనికిని అవగ్రహము (౭) చేత సూచిస్తారు: ఉదాహరణాలు: దశమోఽధ్యాయః, రామోఽర్చ్యుః, పాదోఽన్యేహాభవాత్పునః
 తెలుగు ఇస్కీలోగాని, యూనికోడులగాని అవగ్రహములేనందున పై పదములను వ్రాయలేము.

2. ఆర్ధవిసర్గల లోపం

విసర్గపూర్వకమైన క,ఖలకు జిహ్వమూలీయములని, విసర్గపూర్వకమైన ప,ఫలకు ఉపధ్మానీయములనియు పేరులు.

ఉదాహరణములు : కః కపిః, యః ఖలః, కః కిమ్, కః ఫేలతి

కః పఠతి, ద్రుః ఫలతి, ద్యౌర్నః పితా

ఇటువంటి సందర్భములందు ఆర్ధవిసర్గ వికల్పముగా వచ్చును. కాని ఇస్కీలోగాని, యూనికోడులోగాని ఆర్ధవిసర్గ లేనందున ఈ ఆర్ధవిసర్గను చూపలేము. ముఖ్యంగా వేదగ్రంథాలను వ్రాసేటప్పుడు ఈ లోపం కొట్టవచ్చినట్లుగా కనిపిస్తుంది.

3. వేదస్వరాల లోపం

వేదంలో స్వరాలను గుర్తించడానికి ఉదాత్త, ద్విరుదాత్త, అనుదాత్తసంజ్ఞలను వాడటం పరిపాటి. కాని ఆ చిహ్నాలు తెలుగు యూనికోడులో లేనందువల్ల, వేదాలను తెలుగులో వ్రాయలేకపోతున్నాము. ఉదాహరణకు:

*పురుషసంమితో యజ్ఞో యజ్ఞః పురుషసంమితః| అగ్నే తదస్య కల్పయ త్వగంహి వేత్తే యథా తథగ్ం స్వాహా|
యత్సాకత్రా మనసా దీనదక్షాన యజ్ఞస్య మన్వతే మర్తా సః| అగ్నిష్ట ధోతా క్రతువిద్యోజానన్యజిష్ఠో దేవాగం ఋతుశో
యజాతి స్వాహా|*

4. గ్ం, గ్ం వ్రాయలేకుండటం

పై వేదమంత్రములలో హలంతముతో కూడిన గ, గ్లకు తర్వాత అనుస్వారము (సున్న) వచ్చినది. యజుర్వేదప్రాతిశాఖ్యసూత్రముల ప్రకారము రకారము, శషసహలకు ముందుగల మకారము అనుస్వారముగా (సున్నగా) మాఱును. అట్లు జరిగినప్పుడు ఆ సున్నా అర్ధగకారముకై మగును. కాని యూనికోడు అక్షరనిర్మాణనియమప్రకారం పొల్లు తర్వాత సున్న వచ్చుటకు వీలులేదు. యూనికోడులోని ఈ నిర్బంధమును, కనీసము అర్ధగకారమునకైనా సడలించవలెను.

5. రెండు రకాల అనుస్వారాలు

వాజసానేయిప్రాతిశాఖ్యము ప్రకారము అనుస్వారముతో గూడిన అక్షరమునకు రెండు మాత్రల ప్రమాణమున్నా, అనుస్వారమునకు ముందు ప్రాస్వాచ్చుతో కూడిన వర్ణమున్నపుడు ఆ అనుస్వారప్రమాణము ఒకటిన్నరమాత్రలును, దీర్ఘాచ్చుతో కూడిన వర్ణమున్నపుడు అర్ధమాత్రకు సమానముగను ఉండును. పూర్వపు వేదప్రతులలో ఈ రెండురకములైన అనుస్వారములకు వేర్వేరు చిహ్నము లుండేవి. కాని ఈ కాలమున సామాన్యముగా ఒకే చిహ్నమును (ం) వాడుచున్నారు. తౌకికభాషలో ఒకే చిహ్నమును వాడినను, వేదములను వ్రాయునప్పుడు ఈ అనుస్వారద్వైవిధ్యమును గుర్తించుట యుక్తముగా నుండునని నేను తలంచుచున్నాను.

6. దంత్య చ, జల లోపం

దేవనాగరియూనికోడులో దంత్యజకారమును గుర్తించుటకు "నుక్త" అను చిహ్నమున్నది. కాని ఇట్టిది తెలుగుయూనికోడులో లేనందువల్ల దంత్య చ,జలను, వానితో వచ్చే గుణింతరూపములను వ్రాయలేకున్నాము. నిఘంటువుల నిర్మాణ

ములో సీలోపము కొట్టవచ్చినట్లు కనిపించుచున్నది. అంతేకాక ఇవిలేనిది అక్షరక్రమములో తెలుగుపదములను ఆమరించుట లోపభూయిష్టముగా నుండగలదు.

7. తలకట్టు గుర్తు లోపం

శ్రీ కర్నాటి మల్లేశంగారు నా దృష్టికితెచ్చిన లోపమిది. యూనికోడులో ఇతరాచ్చులను గుర్తించుటకు గుర్తులున్నవిగాని, "అ"కారమును గుర్తించుటకు "తలకట్టు" గుర్తు లేదు. ఈలోపమును పూరించవలసియున్నది.

8. గురు,లఘుసంజ్ఞల లోపం

ఛందోగ్రంథములలో గురు,లఘుక్షరములను ఆ అక్షరములపై నుంచిన అర్థచంద్ర, ఊర్ధ్వరేఖాసంజ్ఞలతో సూచింతురు. ఇచ్చట ఊర్ధ్వరేఖకు వేదమునందలి ఉదాత్తసంజ్ఞను స్వీకరించినను, అర్థచంద్రరేఖ ఇంకను లోపించుచున్నది.

| U U U U |

ఉదాహరణము: **యమాతారాజభానస**

9. సంగీతస్వరసంకేతాల లోపం

త్యాగరాజాది సంగీతకృతులను వ్రాయునప్పుడు గాయకుల సౌకర్యమునకై, వానికి స్వరసంకేతములను కూర్చవలెను. నాకు సంగీతజ్ఞానము శూన్యమైనందున, మిత్రులు శ్రీ శిష్టా శ్రీరామచంద్రమూర్తిగారు సూచించిన కర్ణాటక సంగీతస్వరసంకేతముల సీక్రింద పేర్కొనుచున్నాను. ఒకవేళ హిందుస్తానిసంగీతస్వరసంకేతమునకు అదనపు చిహ్నము లవసరమేమో? నాకు వానిని గుఱించి తెలియదు.

- మంద్రస్థాయిస్వరసంకేతము - స్వరము క్రింది చుక్క

ఉదాహరణము: **స రి గ ఘ ణ డ ని**

- తారస్థాయిస్వరసంకేతము - స్వరము పైని చుక్క

ఉదాహరణము: **స̇ రి̇ గ̇ ఘ̇ ణ̇ డ̇ ని̇**

- ద్విగుణవేగసంకేతము - క్రింది గీటు

ఉదాహరణము: **స రి గ రి గా -**

ఇచ్చట **స రి గ రి** అను స్వరములను రెండింతల వేగముతో పలుకవలెను.

- త్రిగుణవేగసంకేతము - క్రింది రెండు గీట్లు

ఉదాహరణము: **స రి గ రి స రి గ రి గా**

ఇచ్చట **స రి గ రి స రి** అను స్వరములను మూడింతల వేగముతో పలుకవలెను.

- తాళాంగసంపూర్తిసంకేతము - ఒక నిలుపు గీత

ఉదాహరణము: **సరిగరి గపథప**

- తాళావృత్తిసంపూర్తిసంకేతము - రెండు నిలుపు గీతలు

ఉదాహరణము: **సరిగరి గారిగ పగపా దపథా||**

- షడ్జమ-పంచమావరోహణసంకేతము - ఒక వాలుగీత (slash)

ఉదాహరణము: స / ప

- నిషాద-మధ్యమావతరణసంకేతము - ఒక తిర్యగ్స్వక్రరేఖ (backslash)

ఉదాహరణము: ని \ మ

- వరయక్సంకేతము - మీది tilda

ఉదాహరణము: స̃

- తాళసంకేతార్ధవంద్రసంజ్ఞ - మీది tilda

- తాళసంకేతసంజ్ఞలు - ౦|౮

సంగీతమున తాళమును ౦|౮ అను గుర్తులతో సూచింతురు.

ఉదాహరణము: |౦|| (ధ్రువతాళము); |౮౦ (జంపెతాళము)

ఇందులో నిలువురేఖకు పైన పేర్కొన్న తాళాంగసంపూర్తిరేఖను, వలయమునకు తెలుగు సున్నాను వాడుకొన్నను, ఆర్ధవంద్రాకారరేఖ ఇంకను లోపించుచున్నది.

నా దృష్టికి వచ్చినట్టి తెలుగుపుస్తకప్రచురణకు కావలసినవీ, యూనికోడులో(ఇస్కీలో) ప్రస్తుతము లోపించినవీ ఐన అక్షరముల (చిహ్నములను) పైన సూచించినాను. నాకు తెలియనివీ, ఇతరులకు తెలిసినవీ ఐన చిహ్నములు మఱి కొన్ని ఉండవచ్చును. వీటి నన్నిటిని జతగూర్చి, యూనికోడు ప్రమాణపట్టికలో నిర్ధరింపజేయవలసిన అవసరమెంతగానో ఉన్నది. పైన సూచించిన వర్ణములలో, దంత్య చ,జలు, తలకట్టు దక్క, మిగిలినవన్నియు సమస్తభారతీయభాషలకును సామాన్యముగ వర్తించునవే. అందుచేత వీని నన్నిటిని ఒక అనుబంధభారతీయభాషాసంజ్ఞల పట్టికలో చేర్చుట సమంజసముగా నుండును. అప్పుడు ప్రతిభాషలోను వీటిని వేరువేరుగా చేర్చవలసిన అవసరముండదు. ఇట్లు సమస్తభారతీయభాషలకు వర్తించే అనుబంధ పట్టిక నొకదానిని ప్రమాణీకరించి, అందులోని వర్ణములను ఆయాభాషలకు కావలసిన రీతిగా వివిధరూపములలో ప్రదర్శింపవలసినదిగా నొక నిబంధనను యూనికోడులో చేర్చవలసియున్నది. ఇట్టిపట్టిక యూరపియనుభాషల కిదివఱకే యూనికోడులో ప్రమాణీకరింపబడియున్నది. ఇందులో భారతీయభాషలకు కావలసిన వర్ణచిహ్నములు కొన్నియున్నను, వీని నీ భాషలయందు వాడుకొనుటకు ప్రస్తుతము వీలులేకున్నది. అందుచేత పైన సూచించినటువంటి వర్ణముల నన్నిటిని చేర్చి భారతీయభాషలన్నటికీ సర్వసామాన్యమైన ఒక అనుబంధపట్టికను యూనికోడులో నిర్ధరింపవలసిన అవసరమెంతేని యున్నది. ఈ విషయమును యూనికోడుసంఘము దృష్టికి తెచ్చి, ఇది అమలుజరుగునట్లు చూచుట మన కనీసభాధ్యతయని నేను మనవి చేయుచున్నాను. ఇట్టి పట్టికలో భారతీయజ్యోతిశాస్త్రమునకు చెందిన గ్రహాదిచిహ్నములను గూడ చేర్చవచ్చును.