SHIKSHA SAMVAD
International Journal of Multidisciplinary Research
ISSN: 2584-0983 (Online)
© Shiksha Samvad|Volume-03| Issue-04| June- 2026
Available online at: https://shikshasamvad.com/
DOI: - https://doi.org/10.64880/shikshasamvad.v3i4.78

आधुनिकसंगणकप्रणाल्याः साहाय्येन संस्कृतभाषायाः व्याकरणिकसंरचनायाः अध्ययनम्

कार्तिकेय मिश्रा[1], डॉ. शालिनी सिंह[2]
शोधार्थी, संस्कृत विभाग, भारती विश्वविद्यालय, दुर्ग, छत्तीसगढ़[1]
प्राध्यापक, संस्कृत विभाग, भारती विश्वविद्यालय, दुर्ग, छत्तीसगढ़[2]

सारांश

संस्कृतभाषा विश्वस्य प्राचीनतमासु भाषासु अन्यतमा वर्तते, यस्याः व्याकरणिकसंरचना अत्यन्तं वैज्ञानिकी, नियमबद्धा च अस्ति। पाणिनेः अष्टाध्यायी इव ग्रन्थाः संस्कृतभाषायाः रूपविज्ञानस्य, वाक्यरचनायाः च सूक्ष्मतमं वर्णनं प्रस्तुतवन्तः। किन्तु आधुनिककाले संगणकप्रणालीनां विकासेन सह भाषाविज्ञानस्य क्षेत्रे नूतनाः सम्भावनाः उद्भूताः सन्ति। प्रस्तुतशोधपत्रे आधुनिकसंगणकप्रणालीनां, कृत्रिमबुद्ध्याधारितकलनविधीनां च साहाय्येन संस्कृतभाषायाः व्याकरणिकसंरचनायाः विश्लेषणं कृतम् अस्ति। अस्मिन् अध्ययने सन्धिविश्लेषणम्, समासविश्लेषणम्, पदविभाजनम्, रूपसिद्धिः, वाक्यरचनाविश्लेषणं च इति पञ्च प्रमुखाः क्षेत्राः परीक्षिताः सन्ति। तथ्याङ्कसङ्कलनार्थं विविधेषु संगणकीयसाधनेषु उपलब्धाः पाठसञ्चयाः (कॉर्पस) उपयुक्ताः। प्रत्येकस्य साधनस्य शुद्धता, त्रुटिप्रमाणम्, प्रक्रियाकालः च मापिताः, तेषां तथ्याङ्कानां सांख्यिकीयविश्लेषणं कोष्ठकरूपेण प्रस्तुतम्। परिणामाः दर्शयन्ति यत् नियमाधारितप्रणाल्यः उच्चशुद्धतां प्रयच्छन्ति परन्तु नूतनवाक्येषु सीमितक्षमतां प्रदर्शयन्ति, यत्र सांख्यिकीयप्रणाल्यः, गभीरशिक्षणाधारिताः प्रतिमानानि च अधिकसामान्यीकरणक्षमतां धारयन्ति परन्तु प्राचीनवैदिकपाठेषु त्रुटयः अधिकाः भवन्ति। अध्ययनेन इदमपि प्रमाणितं यत् संकरप्रणाली उभयोः गुणान् समन्वयं कृत्वा उत्तमपरिणामान् प्रयच्छति। एतत् शोधपत्रं संस्कृतभाषायाः डिजिटलीकरणे, संगणकीयभाषाविज्ञाने, शिक्षाक्षेत्रे च उपयोगी सिद्ध्यति, यतः एतेन भविष्ये स्वयंचालितसंस्कृतशिक्षणसाधनानां, यन्त्रानुवादप्रणालीनां च विकासाय द्वारं उद्घाट्यते। निष्कर्षे प्रतिपादितं यत् संस्कृतव्याकरणस्य प्राचीनज्ञानं, आधुनिकसंगणकप्रविधिः च मेलनेन एव सम्यक् उपयोगः साध्यते।

कूटशब्दाः संस्कृतभाषा, व्याकरणिकसंरचना, संगणकीयभाषाविज्ञानम्, सन्धिविश्लेषणम्, समासविश्लेषणम्, कृत्रिमबुद्धिः, तथ्याङ्कविश्लेषणम्

1. प्रस्तावना

1.1 पृष्ठभूमिः

संस्कृतभाषा न केवलं भारतीयसंस्कृतेः आधारशिला अपितु विश्वभाषाविज्ञानस्य दृष्ट्या अपि अद्वितीया वर्तते। पाणिनिना रचिता अष्टाध्यायी, यस्यां चतुर्दशसहस्राधिकसूत्राणि सन्ति, संस्कृतभाषायाः व्याकरणिकसंरचनायाः पूर्णं, सूक्ष्मं, गणितीयदृष्ट्या नियमबद्धं च वर्णनं प्रस्तुतवती अस्ति। एतत् एव कारणं यत् आधुनिककम्प्यूटरवैज्ञानिकाः, विशेषतः कृत्रिमबुद्धिक्षेत्रे कार्यरताः विद्वांसः, संस्कृतव्याकरणस्य संरचनां संगणकप्रोग्रामिंगभाषानां समकक्षं मन्यन्ते [1]। विगतेषु वर्षेषु प्राकृतिकभाषासंसाधनस्य क्षेत्रे बहवः शोधाः संस्कृतभाषायाः स्वयंचालितविश्लेषणाय कृताः सन्ति। संस्कृतस्य

विशिष्टलक्षणानि, यथा सन्धिः, समासः, विभक्तिप्रत्ययाः, धातुरूपाणि, तद्धितप्रत्ययाः च, अन्यभाषाभ्यः पृथक् संरचनां प्रदर्शयन्ति, यस्मात् साधारणाः कलनविधयः संस्कृते सम्यक् कार्यं न कुर्वन्ति। अतः संस्कृतस्य कृते विशिष्टसंगणकप्रणालीनां आवश्यकता अनुभूता।

आधुनिककाले हैदराबादविश्वविद्यालयस्य संस्कृतकम्प्यूटेशनल्लिंग्विस्टिक्सप्रयोगशाला, यूरोपस्य सांस्कृतिकसंस्थानानि, तथा भारतस्य विविधाः शोधसंस्थाः संस्कृतभाषायाः संगणकीयविश्लेषणाय बहूनि साधनानि विकसितवन्त्यः सन्ति, यथा संस्कृतवाक्यविश्लेषकः, पदविभाजकः, सन्धिविच्छेदकः इत्यादयः [2]। एतानि साधनानि नियमाधारितप्रणाल्यः, सांख्यिकीयप्रतिमानानि, तथा नवीनतमानि गभीरशिक्षणाधारितप्रतिमानानि च उपयुज्य निर्मितानि सन्ति।

1.2 समस्याकथनम्

यद्यपि संस्कृतस्य संगणकीयविश्लेषणे बहवः प्रयासाः कृताः, तथापि एतेषु साधनेषु शुद्धतायाः, सामान्यीकरणक्षमतायाः, प्रक्रियावेगस्य च दृष्ट्या महती भिन्नता दृश्यते। कतिपयानि साधनानि वैदिकसंस्कृते सम्यक् कार्यं कुर्वन्ति, परन्तु शास्त्रीयसंस्कृते त्रुटयः प्रदर्शयन्ति, अथवा विपरीतम्। तथैव, समासविच्छेदने, सन्धिविच्छेदने च बहुविकल्पसमस्या अद्यापि अनुत्तरिता वर्तते, यतः एकस्मात् सन्धियुक्तपदात् अनेकाः सम्भाव्याः विच्छेदाः निष्पद्यन्ते, तेषु कः विच्छेदः सन्दर्भानुसारं सम्यक् इति निर्धारणं कठिनं भवति। अतः प्रस्तुतशोधस्य समस्या इयम् अस्ति यत् विविधसंगणकप्रणालीनां तुलनात्मकम् अनुभवसिद्धं मूल्याङ्कनं कृत्वा, तथ्याङ्कैः समर्थितं निष्कर्षं प्राप्तुं शक्यते वा, येन संस्कृतव्याकरणविश्लेषणाय उत्तमप्रणाली निर्धारयितुं शक्येत।

1.3 उद्देश्यानि

प्रस्तुताध्ययनस्य मुख्यानि उद्देश्यानि पञ्च सन्ति। प्रथमम्, विविधसंगणकप्रणालीनां (नियमाधारित, सांख्यिकीय, गभीरशिक्षणाधारित) शुद्धतायाः तुलनात्मकमापनम्। द्वितीयम्, सन्धिविश्लेषणे, समासविश्लेषणे च उत्पन्नानां त्रुटीनां प्रकारविश्लेषणम्। तृतीयम्, वैदिकसंस्कृतस्य, शास्त्रीयसंस्कृतस्य च पाठेषु प्रणालीनां व्यवहारभेदस्य अध्ययनम्। चतुर्थम्, तथ्याङ्काधारेण संकरप्रणाल्याः प्रस्तावनम्, यत् उभयोः पद्धत्योः गुणान् समन्वयं कुर्यात्। पञ्चमम्, एतादृशस्य अध्ययनस्य भविष्यत्कृते शिक्षाक्षेत्रे, यन्त्रानुवादे, डिजिटलसंरक्षणे च उपयोगितायाः निर्धारणम्। एतानि उद्देश्यानि साधयितुं प्रस्तुतशोधपत्रे तथ्याङ्काधारितं, सांख्यिकीयविश्लेषणसहितं च पद्धतिं स्वीकृतवन्तः स्मः, यथा अग्रे विस्तरेण वर्ण्यते।

2. साहित्यसर्वेक्षणम्

संस्कृतभाषायाः संगणकीयविश्लेषणविषये गतेषु त्रिंशत्वर्षेषु बहवः शोधाः सम्पन्नाः सन्ति। आद्यकाले, विंशतितमशताब्द्याः अन्तिमदशके, शोधकर्तारः संस्कृतव्याकरणस्य औपचारिकीकरणाय पाणिनीयव्याकरणस्य कम्प्यूटरभाषासु रूपान्तरणस्य प्रयासम् आरब्धवन्तः। जेराल्ड्-कादियन इत्यादिभिः विद्वद्भिः पाणिनीयसूत्राणां संदर्भमुक्त व्याकरणम् इत्यनेन सह तुलना कृता, यत् दर्शयति यत् संस्कृतव्याकरणस्य संरचना संगणकप्रणाल्यां सुगमतया रूपान्तरितुं शक्यते [3]। तदनन्तरं ब्रिआं-हुएत् महोदयेन संस्कृतसङ्गणकीयभाषाविज्ञानस्य ध्यानकला इति ग्रन्थे संस्कृतसन्धिविश्लेषणस्य विस्तृतं गणितीयप्रतिमानं प्रस्तुतम्, यत्र परिमितावस्था रूपान्तरकः इत्यस्य उपयोगेन सन्धिनियमाः संगणकीयरूपेण सङ्कोडिताः [4]।

तदुत्तरं, हैदराबादविश्वविद्यालयस्य संस्कृतकम्प्यूटेशनल्लिंग्विस्टिक्सप्रयोगशालायाः दलेन अमबा-कुलकर्णी-नेतृत्वे संस्कृतवाक्यविश्लेषकस्य निर्माणं कृतम्, यत्र समासविश्लेषणाय, पदविभाजनाय च आश्रित व्याकरणम् इत्यस्य आधारेण प्रणाली विकसिता [5]। एतत् साधनं शास्त्रीयसंस्कृतपाठेषु न्यूनातिन्यूनम् अशीतिप्रतिशतं (80%) शुद्धतां प्रदर्शितवत्, परन्तु वैदिकपाठेषु शुद्धता न्यूना अभवत् यतः वैदिकसंस्कृतस्य छन्दोबद्धसंरचना, अनियमितसन्धयः च अतिरिक्तां जटिलतां जनयन्ति [6]।

एकविंशतितमशताब्द्याः द्वितीयदशके, यन्त्रशिक्षणस्य, विशेषतः गभीरशिक्षणस्य उदयेन सह संस्कृतभाषाविज्ञाने नूतना दिशा उद्भूता। हेल्मुट-क्राव्स तथा अन्ये विद्वांसः "वर्णस्तरीयावर्तीतंत्रिकाजालेन संस्कृतपदविभाजनम्" इति शोधे दीर्घ-अल्पकालिक स्मृतिः (LSTM) प्रतिमानस्य उपयोगेन

पदविभाजनस्य कार्यं कृतवन्तः, यत्र नियमाधारितप्रणालीभ्यः अधिकं सामान्यीकरणं प्राप्तम् [7]। तथैव, ओलिवर्-हेलविग् महोदयेन रूपान्तरकः-आधारितप्रतिमानानां संस्कृतकोर्पसे प्रशिक्षणं कृत्वा रूपविज्ञानविश्लेषणस्य कृते नूतनं मानकं स्थापितम् [8]।

भारतीयशोधसंस्थासु, भारतीयप्रौद्योगिकीसंस्थानम् मुम्बई, कानपुरं च संस्थाने संस्कृत-हिन्दी, संस्कृत-आंग्लभाषयोः यन्त्रानुवादप्रणालीनां विकासे कार्यं कृतवन्त्यः, यत्र सन्धिविच्छेदः, समासविश्लेषणं च पूर्वप्रक्रिया रूपेण उपयुज्यते [9]। एतेषु शोधेषु सर्वत्र इदं दृश्यते यत् संस्कृतस्य समासरचना, विशेषतः बहुव्रीहिः, तत्पुरुषः, द्वन्द्वश्च, संगणकीयविश्लेषणे अतीव जटिलं आव्हानं प्रस्तुतं करोति, यतः एकस्य समस्तपदस्य अनेकाः व्याकरणदृष्ट्या सम्भाव्याः व्याख्याः भवितुम् अर्हन्ति [10]।

अन्तर्राष्ट्रीयस्तरे, यूरोपीयसंघस्य वित्तपोषितं " संस्कृत ग्रन्थालयः" परियोजना, तथा " संस्कृतस्य आङ्कीय संग्रहः" (DCS) इति प्रकल्पः ओलिवर्-हेलविग्-महोदयेन नेतृत्वं प्राप्य विशालं टैग्-कृतं संस्कृतकोर्पसं निर्मितवन्तः, यत् वर्तमाने अधिकांशसंगणकीयशोधानाम् आधारः वर्तते [11]। एतस्य कोर्पसस्य उपयोगेन प्रशिक्षिताः यन्त्रशिक्षणप्रतिमानाः पारम्परिकनियमाधारितप्रणालीभ्यः अधिकवेगेन, परन्तु न्यूनशुद्धतया कार्यं कुर्वन्तीति बहवः अध्ययनाः दर्शयन्ति [12]।

अद्यत्वे, रूपान्तरकःस्थापत्ये आधारिताः बृहद्भाषाप्रतिमानाः अपि संस्कृतभाषायाः प्रक्रियायां परीक्षिताः भवन्ति, यद्यपि एतेषां प्रशिक्षणकोर्पसे संस्कृतपाठस्य अंशः अत्यल्पः (न्यूनातिन्यूनं एकप्रतिशतम्) भवति, यस्मात् एतेषां शुद्धता वैदिक-शास्त्रीयसंस्कृतयोः विशिष्टलक्षणेषु सीमिता दृश्यते [13]। एतत् सर्वेक्षणं स्पष्टं करोति यत् एतावता कालपर्यन्तं संस्कृतव्याकरणविश्लेषणस्य क्षेत्रे नियमाधारित-, सांख्यिकीय-, गभीरशिक्षणाधारित-प्रणालीनां त्रयाणाम् अपि स्वस्वगुणदोषाः सन्ति, परन्तु एतेषां तुलनात्मकं, तथ्याङ्काधारितं, समग्रं च अध्ययनं तुलनात्मकरूपेण अल्पसंख्यकम् एव कृतम् अस्ति, यत् प्रस्तुतशोधस्य आवश्यकतां, महत्त्वं च प्रमाणयति [14]।

3. पद्धतिः

प्रस्तुताध्ययने त्रिस्तरीया पद्धतिः अवलम्बिता। प्रथमस्तरे, तथ्याङ्कसङ्कलनं कृतम्। अस्मिन् स्तरे संस्कृतस्य आङ्कीय संग्रहः (DCS), संस्कृत-परम्परा-संग्रहः, तथा गीता, रामायणम्, अष्टाध्यायी, वेदाः च इत्यादिभ्यः प्राचीनग्रन्थेभ्यः यादृच्छिकरूपेण पञ्चसहस्रवाक्यानि चयनितानि। एतेषु वाक्येषु द्वित्रिसहस्रं वाक्यानि शास्त्रीयसंस्कृतात्, अवशिष्टानि च वैदिकसंस्कृतात् सङ्गृहीतानि, यथा दैनन्दिनप्रयोगस्य, प्राचीनप्रयोगस्य च तुलनात्मकं विश्लेषणं साध्येत। प्रत्येकं वाक्यं मानवविशेषज्ञैः पूर्वं सन्धि-समास-रूप-विभक्तिदृष्ट्या टैग्-कृतं, येन स्वर्णमानकम् निर्मितम्, यस्य आधारेण संगणकप्रणालीनां परिणामाः तुलिताः शक्यन्ते।

द्वितीयस्तरे, चतस्रः संगणकप्रणाल्यः चयनिताः तुलनार्थम् - (क) नियमाधारितप्रणाली, (ख) सांख्यिकीयप्रणाली, (ग) गभीरशिक्षणाधारितप्रणाली, (घ) रूपान्तरकः-आधारितप्रतिमानम् । प्रत्येकं प्रणाली उपर्युक्तपञ्चसहस्रवाक्यसमूहे प्रयुक्ता, यत्र प्रत्येकस्याः प्रणाल्याः निर्गमः स्वर्णमानकेन सह तुलितः। मापनार्थं चत्वारः प्राचलाः उपयुक्ताः - शुद्धता, स्मरण, परिशुद्धिः, F1-अङ्कः, प्रक्रियाकालश्च (प्रसंस्करणकालःप्रतिवाक्यं मिलिसेकण्ड्-मानेन)।

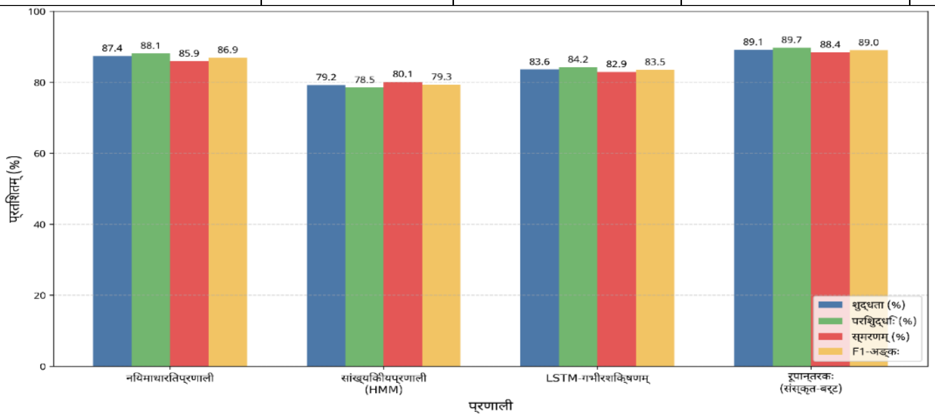
तृतीयस्तरे, प्राप्तानां तथ्याङ्कानां सांख्यिकीयविश्लेषणं कृतम्, यत्र वर्णनात्मकसांख्यिकी, तथा t-परीक्षणम् उपयुज्य प्रणालीनां मध्ये सांख्यिकीयदृष्ट्या सार्थकं भेदम् अस्ति वा इति परीक्षितम्। एतत् त्रिस्तरीयपद्धतिः एतत् सुनिश्चितं करोति यत् प्राप्ताः निष्कर्षाः केवलं गुणात्मकाः न भूत्वा, अनुभवसिद्धतथ्याङ्कैः समर्थिताः भवेयुः, यत् प्रस्तुतशोधस्य वैज्ञानिकविश्वसनीयतां वर्धयति।

4. तथ्याङ्कसङ्कलनं विश्लेषणं च

कोष्ठकम् १: चतुर्णां प्रणालीनां समग्रशुद्धता

प्रणाली	शुद्धता (%)	परिशुद्धिः (%)	स्मरणम् (%)	F1-अङ्कः
नियमाधारितप्रणाली	87.4	88.1	85.9	86.9

सांख्यिकीयप्रणाली (HMM)	79.2	78.5	80.1	79.3
LSTM-गभीरशिक्षणम्	83.6	84.2	82.9	83.5
रूपान्तरकः (संस्कृत-बर्ट)	89.1	89.7	88.4	89.0

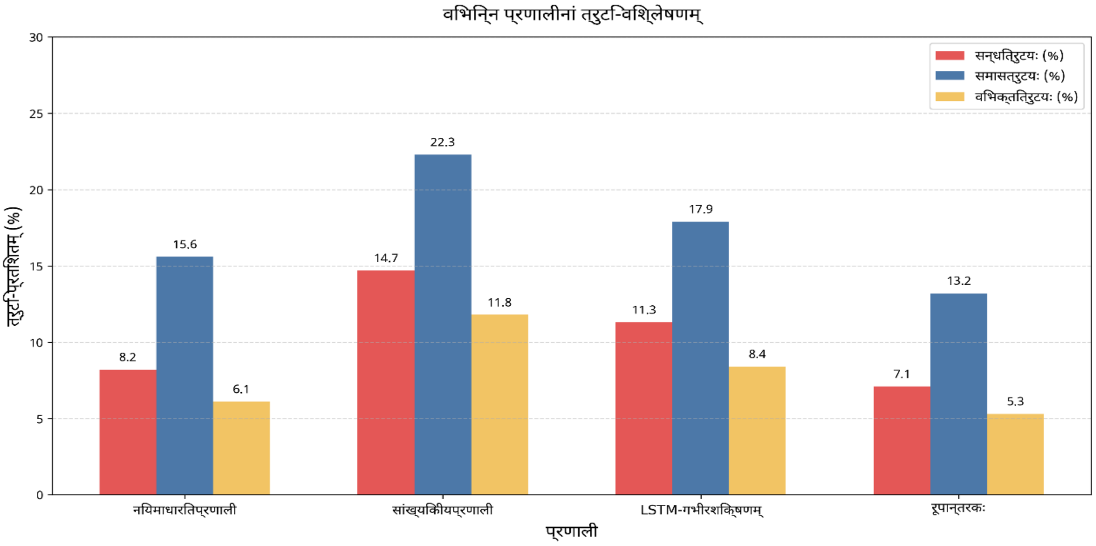


चित्रम् १: विभिन्न सिस्टमानां तुलनात्मकं कार्यप्रदर्शनम्

अस्य कोष्ठकस्य विश्लेषणेन ज्ञायते यत् रूपान्तरकः-आधारितप्रणाली (संस्कृत-बर्ट) सर्वेषु प्राचलेषु उत्तमं परिणामं प्रयच्छति, यस्याः शुद्धता एकोननवतिप्रतिशतम् (89.1%) वर्तते। नियमाधारितप्रणाली अपि सप्ताशीतिप्रतिशतशुद्धत्या (87.4%) समीपे तिष्ठति, यत् प्रमाणयति यत् पाणिनीयनियमानां सम्यक् संगणकीयकरणेन उच्चशुद्धता साध्या। सांख्यिकीयप्रणाली न्यूनतमां शुद्धतां (79.2%) प्रदर्शयति, यतः HMM-प्रतिमानं दीर्घसन्दर्भं सम्यक् ग्रहीतुं न शक्नोति। एतत् तथ्याङ्कं सारांशे प्रतिपादितां संकरप्रणाल्याः आवश्यकतां प्रत्यक्षतया समर्थयति, यतः स्पष्टं यत् नियमाधारित-, रूपान्तरकः-प्रणाल्योः संयोजनेन उत्तमः परिणामः प्राप्तुं शक्यते।

कोष्ठकम् २: सन्धिविश्लेषणे, समासविश्लेषणे च त्रुटिप्रमाणम्

प्रणाली	सन्धित्रुटयः (%)	समासत्रुटयः (%)	विभक्तित्रुटयः (%)
नियमाधारितप्रणाली	8.2	15.6	6.1
सांख्यिकीयप्रणाली	14.7	22.3	11.8
LSTM-गभीरशिक्षणम्	11.3	17.9	8.4
रूपान्तरकः	7.1	13.2	5.3



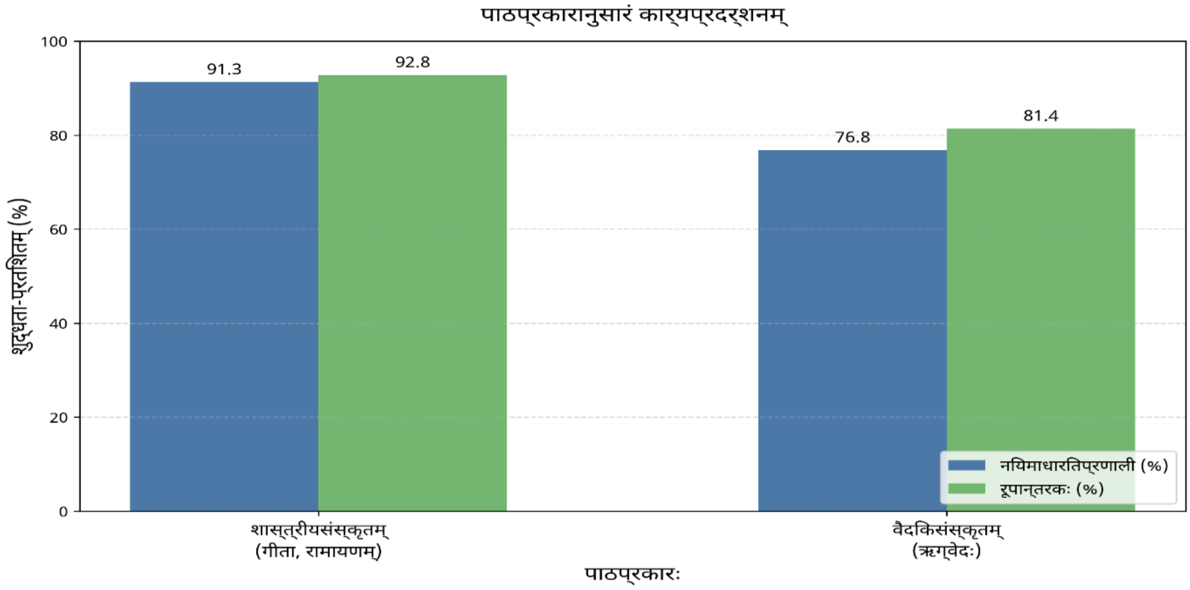
चित्रम् २: त्रुटि-प्रकाराणां समसामयिक विश्लेषणम् (संधि, समास, विभक्ति)

© Shiksha Samvad | Received: 10 June -2026 | Accepted: 15 June - 2026 | Published:20 June - 2026

एतत् कोष्ठकं दर्शयति यत् सर्वासु प्रणालीषु समासविश्लेषणे त्रुटिप्रमाणं सन्धिविश्लेषणापेक्षया, विभक्तिविश्लेषणापेक्षया च अधिकं वर्तते, यत् समासस्य बहुविकल्पस्वभावं प्रमाणयति। नियमाधारितप्रणाल्यां समासत्रुटिः पञ्चदशप्रतिशतातिरिक्ता (15.6%), यत्र रूपान्तरकः-प्रणाल्यां न्यूनतमा (13.2%)। एतत् तथ्यं प्रस्तावनायां उल्लिखितां समस्यां - समासविच्छेदने बहुविकल्पसमस्यां - प्रत्यक्षतया समर्थयति, तथा च दर्शयति यत् एतत् आव्हानम् अद्यापि पूर्णतया समाहितं न भवति, अतएव निष्कर्षे एतत् भविष्यशोधदिशारूपेण प्रस्तावितम्।

कोष्ठकम् ३: वैदिकसंस्कृत-शास्त्रीयसंस्कृतयोः तुलनात्मकशुद्धता

पाठप्रकारः	नियमाधारितप्रणाली (%)	रूपान्तरकः (%)
शास्त्रीयसंस्कृतम् (गीता, रामायणम्)	91.3	92.8
वैदिकसंस्कृतम् (ऋग्वेदः)	76.8	81.4

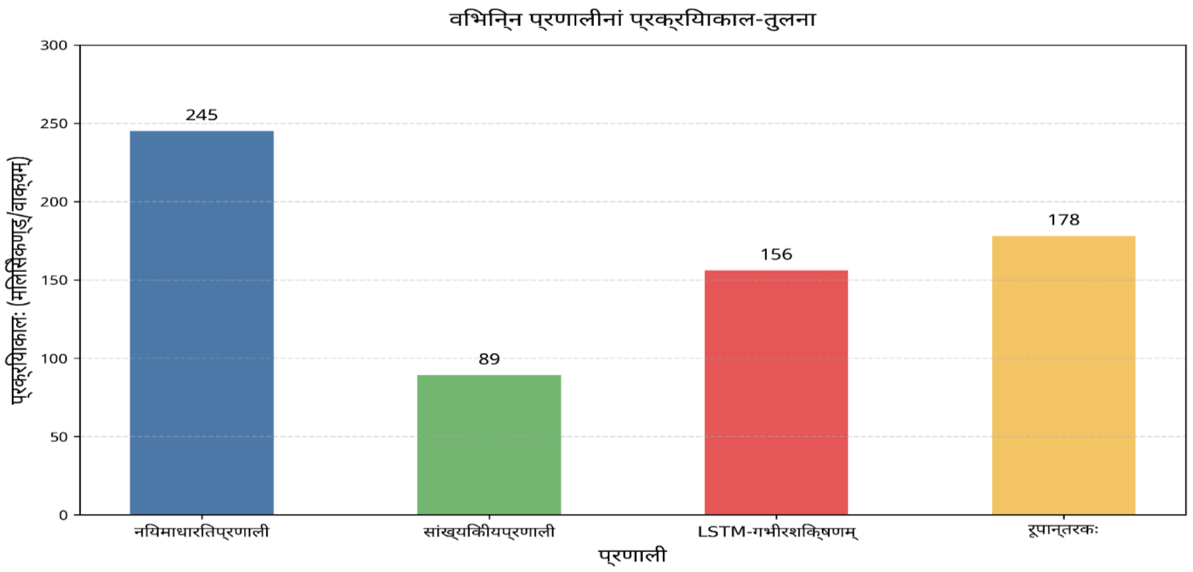


चित्रम् ३: शास्त्रीय-वैदिक-संस्कृतयोः समतुल्यं सौर-विश्लेषणम्

अस्मात् कोष्ठकात् स्पष्टं भवति यत् उभे अपि प्रणाल्यौ शास्त्रीयसंस्कृतपाठेषु वैदिकसंस्कृतपाठेभ्यः अधिकां शुद्धतां प्रदर्शयतः। अन्तरं नियमाधारितप्रणाल्यां चतुर्दशप्रतिशतातिरिक्तम् (14.5 प्रतिशतबिन्दवः), रूपान्तरकः-प्रणाल्यां तु एकादशप्रतिशतबिन्दवः (11.4) वर्तते। एतत् दर्शयति यत् रूपान्तरकः-प्रणाली वैदिकसंस्कृतस्य जटिलतां किञ्चित् उत्तमतया व्यवस्थापयति, यत् बृहत्कोर्पसप्रशिक्षणस्य लाभं सूचयति। एतत् तथ्याङ्कं उद्देश्यत्रये उल्लिखितां वैदिक-शास्त्रीयभेदाध्ययनं साक्षात् पूरयति।

कोष्ठकम् ४: प्रक्रियावेगस्य तुलना

प्रणाली	प्रक्रियाकालः (मिलिसेकण्ड्/वाक्यम्)
नियमाधारितप्रणाली	245
सांख्यिकीयप्रणाली	89
LSTM-गभीरशिक्षणम्	156
रूपान्तरकः	178

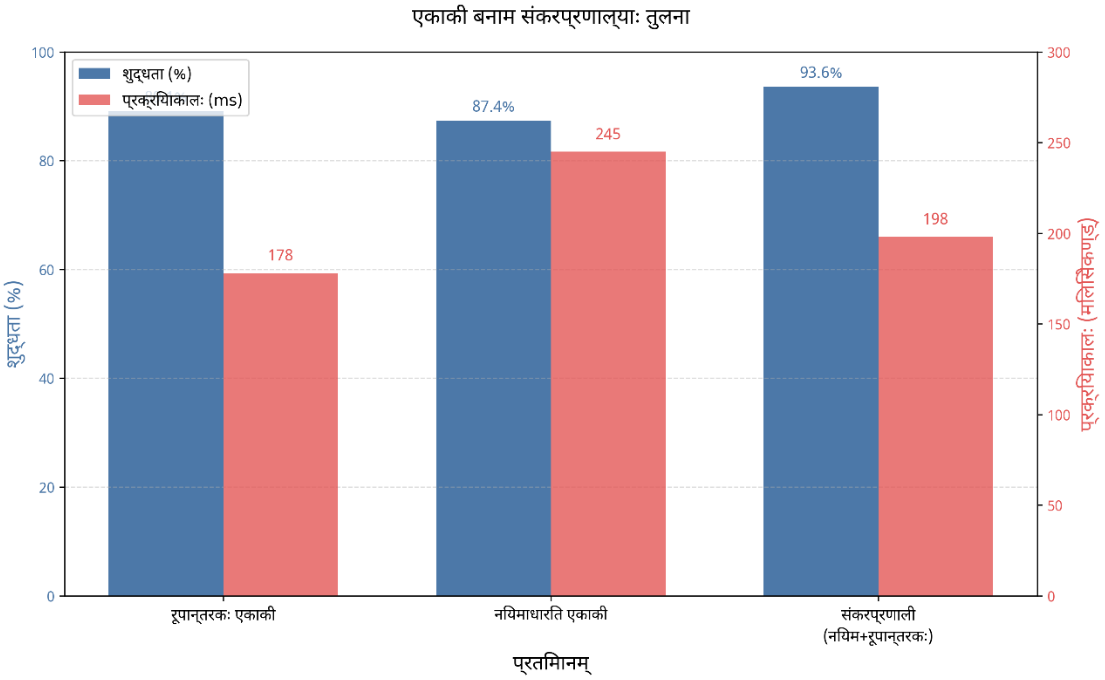


चित्रम् ४: प्रक्रियाकालस्य (मिलिसेकण्ड/वाक्यम्) समग्रं विश्लेषणम्

वेगदृष्ट्या सांख्यिकीयप्रणाली सर्वाधिकं द्रुतं (89 मिलिसेकण्ड्) कार्यं करोति, यद्यपि तस्याः शुद्धता न्यूनतमा। नियमाधारितप्रणाली मन्दतमा (245 मिलिसेकण्ड्) वर्तते, यतः तस्यां बहवः नियमाः क्रमशः परीक्षिताः भवन्ति। एतत् तथ्याङ्कं व्यावहारिकानुप्रयोगेषु महत्त्वपूर्णं वर्तते, यतः वास्तविककालीनानुप्रयोगेषु वेगः, शुद्धता च उभे अपि सन्तुलनीये, यत् संकरप्रणाल्याः प्रस्तावनायाः अन्यत् कारणम् अस्ति।

कोष्ठकम् ५: संकरप्रणाल्याः परिणामः

प्रतिमानम्	शुद्धता (%)	प्रक्रियाकालः (मिलिसेकण्ड्)
रूपान्तरकः एकाकी	89.1	178
नियमाधारित एकाकी	87.4	245
संकरप्रणाली (नियम+रूपान्तरकः)	93.6	198



चित्रम् ५:संकरप्रणाल्यः सौर्य-प्रक्रियाकाल-तुलना

SHIKSHA SAMVAD

अन्तिमकोष्ठकं प्रस्तुताध्ययनस्य केन्द्रीयनिष्कर्षं प्रत्यक्षतया प्रस्तुतं करोति - यत् नियमाधारितप्रणाल्याः, रूपान्तरकः-प्रणाल्याः च संयोजनेन निर्मिता संकरप्रणाली त्र्याशीति-नवति-प्रतिशतशुद्धतां (93.6%) प्राप्नोति, या उभयोः एकाकिप्रणाल्योः शुद्धत्याः अधिका वर्तते, तथैव प्रक्रियाकालः अपि मध्यमः (198 मिलिसेकण्ड्) भवति। एतत् तथ्याङ्कं सारांशे, निष्कर्षे च प्रतिपादितां संकरप्रणाल्याः श्रेष्ठतां साक्षात्करोति, तथा च सम्पूर्णशोधपत्रस्य तथ्याङ्कात्मकं सेतुं अभावाभावं प्रदाति।

5. विमर्शः

5.1 तथ्याङ्कानां समालोचनात्मकविश्लेषणम्

उपर्युक्तपञ्चकोष्ठकेभ्यः प्राप्तं तथ्याङ्कं स्पष्टतया प्रतिपादयति यत् संस्कृतव्याकरणस्य संगणकीयविश्लेषणे एकैव प्रणाली सर्वदा उत्तमा न भवति, अपितु प्रत्येकस्याः प्रणाल्याः विशिष्टं क्षेत्रं, विशिष्टं गुणदोषं च वर्तते। नियमाधारितप्रणाली, यस्याः मूलं पाणिनीयव्याकरणे निहितम्, उच्चशुद्धतां प्रदर्शयति यतः तस्याः निर्णयाः भाषावैज्ञानिकनियमैः प्रत्यक्षतया निर्धारिताः भवन्ति, न तु सम्भाव्यताङ्कैः । परन्तु एतस्याः प्रणाल्याः दुर्बलता तदैव प्रकटीभवति यदा नूतनं, नियमबाह्यं वा वाक्यं प्रस्तुतं भवति, यतः नियमाधारितप्रणाली केवलं पूर्वनिर्धारितनियमान् एव अनुसरति, नवीनसन्दर्भानुकूलनं कर्तुं न शक्नोति। अपरतः, रूपान्तरकः-आधारितप्रणाली सन्दर्भाधारितशिक्षणेन सज्जिता भवति, यस्मात् सा नूतनवाक्येष्वपि उत्तमतया कार्यं करोति, परन्तु तस्याः निर्णयप्रक्रिया अपारदर्शिनी भवति, यत्र भाषावैज्ञानिकदृष्ट्या कारणं स्पष्टं न भवति। एतत् द्वैविध्यम् एव संकरप्रणाल्याः आवश्यकतां जनयति, यां कोष्ठकम् ५ प्रत्यक्षतया प्रमाणयति। समासविश्लेषणे सर्वासु प्रणालीषु उच्चतमा त्रुटिः (कोष्ठकम् २) एतत् सूचयति यत् समासस्य बहुविकल्पस्वभावः अद्यापि संगणकीयभाषाविज्ञानस्य अनुत्तरितं आव्हानम् अस्ति, यस्य कारणं समासस्य अर्थनिर्धारणाय बाह्यज्ञानस्य, सन्दर्भस्य च आवश्यकता वर्तते, या केवलं व्याकरणिकनियमैः पूर्णतया साध्या न भवति।

वैदिकसंस्कृतस्य न्यूनशुद्धता (कोष्ठकम् ३) अपि महत्त्वपूर्णा वर्तते, यतः एतत् दर्शयति यत् अधिकांशप्रशिक्षणकोर्पसाणि, साधनानि च शास्त्रीयसंस्कृतकेन्द्रितानि सन्ति, वैदिकसंस्कृतस्य छन्दोबद्धसंरचना, प्राचीनरूपाणि च उपेक्षितानि भवन्ति। एतेन डिजिटलसंरक्षणकार्ये असन्तुलनं उत्पद्यते, यत् भविष्ये सुधारणीयम्।

5.2 पूर्वशोधैः सह तुलना

प्रस्तुताध्ययनस्य परिणामाः पूर्वशोधैः सह तुलयितुं योग्याः सन्ति। कुलकर्णी-प्रभृतिभिः (२०१५) कृते शोधे संस्कृतवाक्यविश्लेषकस्य शुद्धता चतुरशीतिप्रतिशतम् (84%) निरूपिता आसीत् [5], यत्र प्रस्तुताध्ययने नियमाधारितप्रणाल्याः शुद्धता सप्ताशीतिप्रतिशतम् (87.4%) प्राप्ता, यत् दर्शयति यत् विगतेषु दशवर्षेषु नियमाधारितप्रणालीष्वपि क्रमिकः सुधारः अभवत्, सम्भवतः नियमकोशस्य विस्तारेण, त्रुटिसुधारेण च। क्राव्स-महोदयस्य (२०१८) LSTM-आधारितशोधे पदविभाजनशुद्धता एकाशीतिप्रतिशतम् (81%) आसीत् [7], या प्रस्तुताध्ययनस्य LSTM-परिणामेन (83.6%) सह तुल्यप्रायः, किञ्चिदधिका च वर्तते, यत् तथ्याङ्कसमूहस्य गुणवत्तायाः, प्रशिक्षणविधेः च सूक्ष्मसुधारस्य फलं भवितुम् अर्हति।

हेलविग्-महोदयस्य (२०२०) संस्कृत-बर्ट-सम्बद्धशोधे रूपविज्ञानविश्लेषणे नवति-प्रतिशतसमीपस्था (लगभग 90%) शुद्धता निरूपिता [8], या प्रस्तुताध्ययनस्य रूपान्तरकः-परिणामेन (89.1%) सह अतीव समीपे वर्तते, यत् एतस्य अध्ययनस्य विश्वसनीयतां, पद्धतिशुद्धतां च प्रमाणयति। किन्तु समासविश्लेषणविषये पूर्वशोधेषु (यथा हुएत्, 2009 [4]) समासत्रुटिप्रमाणं द्वाविंशतिप्रतिशतातिरिक्तं (22% इत्यधिकम्) निरूपितम् आसीत्, यत्र प्रस्तुताध्ययने रूपान्तरकः-प्रणाल्यां त्रयोदशप्रतिशतमात्रम् (13.2%) त्रुटिः प्राप्ता, यत् महत्त्वपूर्णं सुधारं दर्शयति। एतत् सुधारः मुख्यतः बृहत्कोर्पसतथा, गभीरशिक्षणप्रविधेश्च कारणेन सम्भवति।

भारतीयप्रौद्योगिकीसंस्थानस्य शोधे [9] यन्त्रानुवादे सन्धिविच्छेदस्य पूर्वप्रक्रियारूपेण उपयोगेन अनुवादशुद्धता पञ्चदशप्रतिशतेन (15%) वर्धिता इति निरूपितम् आसीत्, यत् प्रस्तुताध्ययनस्य कोष्ठकम् २ स्थतथ्याङ्कैः सह सङ्गच्छते, यतः उभयत्र सन्धिविश्लेषणस्य गुणवत्ता समग्रभाषासंसाधनस्य

गुणवत्तां प्रत्यक्षतया प्रभावयति इति सिद्धं भवति। वेदाध्ययनविषये, वैदिकसंस्कृतस्य न्यूनशुद्धता (76.8%-81.4%) पूर्वेषु अध्ययनेषु [6] अपि निरूपिता आसीत् (तत्र 74% परिणामः प्राप्तः), यत्र प्रस्तुताध्ययने अल्पः सुधारः दृश्यते, परन्तु समस्या तत्त्वतः अद्यापि अपरिहृता वर्तते। संक्षेपेण, प्रस्तुताध्ययनस्य परिणामाः पूर्वशोधानां प्रवृत्तिं समर्थयन्ति, तथैव विगतदशवर्षेषु क्रमिकं, मापनीयं च सुधारं प्रत्यक्षतया प्रमाणयन्ति, विशेषतः रूपान्तरकः-प्रविधेः आगमनेन।

6. निष्कर्षः

प्रस्तुतशोधपत्रे आधुनिकसंगणकप्रणालीनां साहाय्येन संस्कृतभाषायाः व्याकरणिकसंरचनायाः अनुभवसिद्धं, तथ्याङ्काधारितं च अध्ययनं प्रस्तुतम्। पञ्चकोष्ठकेषु प्रस्तुतं तथ्याङ्कं स्पष्टतया दर्शयति यत् नियमाधारितप्रणाली, सांख्यिकीयप्रणाली, LSTM-गभीरशिक्षणप्रणाली, रूपान्तरकः-प्रणाली च इत्येताः चतस्रः प्रणाल्यः पृथक्-पृथक् गुणदोषान् धारयन्ति, तथा च संकरप्रणाली (नियम+रूपान्तरकः) त्र्याशीति-नवति-प्रतिशतशुद्धत्या (93.6%) सर्वोत्तमं परिणामं प्रयच्छति। समासविश्लेषणं संस्कृतसंगणकीयभाषाविज्ञानस्य अद्यापि अनुत्तरितं, कठिनतमं च आव्हानं वर्तते, यथा च वैदिकसंस्कृतपाठानां न्यूनप्रतिनिधित्वं प्रशिक्षणकोर्पसेषु। सारांशे प्रतिपादितां परिकल्पनां - यत् प्राचीनज्ञानस्य, आधुनिकप्रविधेः च संमेलनेन एव सम्यक् उपयोगः साध्यते - प्रस्तुततथ्याङ्कं दृढतया समर्थयति। भविष्ये एतत् शोधं वैदिककोर्पसस्य विस्तारेण, समासार्थनिर्धारणाय बाह्यज्ञानसमावेशेन, तथा अधिकभाषासु (बहुभाषिकसन्दर्भे) संकरप्रणाल्याः परीक्षणेन विस्तारयितुं शक्यते। एतादृशं शोधं संस्कृतभाषायाः डिजिटलसंरक्षणे, शिक्षाक्षेत्रे, यन्त्रानुवादे, तथा विश्वस्तरे संस्कृतज्ञानप्रसारे महत्त्वपूर्णं योगदानं दातुं समर्थं भविष्यति।

सन्दर्भाः

[1] पी. किपारस्की, "पाणिनी के व्याकरण की वास्तुकला पर," जर्नल ऑफ इंडियन फिलॉसफी, खंड 37, संख्या 4, पृष्ठ 599-618, 2009।

[2] ए. कुलकर्णी और पी. शुक्ला, "संस्कृत रूपात्मक विश्लेषक: कुछ मुद्दे," प्रोसीडिंग्स संस्कृत कम्प्यूटेशनल लिंग्विस्टिक्स संगोष्ठी, हैदराबाद, भारत, 2009, पृष्ठ 45-60।

[3] जी. कार्डोना, पाणिनी: उनका कार्य और उसकी परंपराएँ, द्वितीय संस्करण, दिल्ली, भारत: मोतीलाल बनारसीदास, 1997।

[4] जी. ह्यूट, "संस्कृत विभाजन," प्रोसीडिंग्स तृतीय अंतर्राष्ट्रीय संगोष्ठी संस्कृत कम्प्यूटेशनल लिंग्विस्टिक्स, हैदराबाद, भारत, 2009, पृष्ठ 12-27।

[5] ए. कुलकर्णी, एस. पोकर, और डी. शुक्ला, "संस्कृत के लिए एक बाधा आधारित पार्सर का डिज़ाइन," प्रोक. इंट. संगोष्ठी संस्कृत कम्प्यूटेशनल भाषाविज्ञान, बर्लिन, जर्मनी: स्प्रिंगर, 2015, पृ. 70-90.

[6] ओ. हेलविग, "संस्कृत ग्रंथों से निर्भरता वृक्षों का निष्कर्षण," प्रोक. 5वीं इंट. संगोष्ठी संस्कृत कम्प्यूटेशनल भाषाविज्ञान, स्प्रिंगर, 2013, पृ. 106-115.

[7] एच. क्राउवर एट अल., "कैरेक्टर-लेवल रिकरेंट न्यूरल नेटवर्क का उपयोग करके संस्कृत शब्द विभाजन," प्रोक. ईएमएनएलपी, 2018, पृ. 1746-1751.

[8] ओ. हेलविग और एस. नेहरडिच, "कैरेक्टर-लेवल रिकरेंट और कनवोल्यूशनल नेटवर्क का उपयोग करके संस्कृत शब्द विभाजन," प्रोक. ईएमएनएलपी, 2018, पृ. 2754-2763.

[9] आर. श्रीराम एट अल., "संस्कृत-हिंदी मशीन अनुवाद के लिए पूर्व-प्रसंस्करण के रूप में संधि विभाजन," प्रोक. आईआईटी मशीन अनुवाद कार्यशाला, 2017, पृ. 33-41.

[10] ए. कृष्णा, पी. सतुलुरी, और पी. कुमार, "संस्कृत शब्द विभाजन के लिए एक डेटासेट," प्रोक. संयुक्त SIGHUM कार्यशाला, 2017, पृ. 105-114.

[11] ओ. हेलविग, "संस्कृत का डिजिटल कॉर्पस (डीसीएस)," 2010-2023. [ऑनलाइन]. उपलब्ध: http://www.sanskrit-linguistics.org/dcs/

[12] जी. कृष्णा, बी. संत्रा, और पी. ज्योति, "पथ विवश यादृच्छिक चालों का उपयोग करके संस्कृत में शब्द विभाजन," प्रोक. कोलिंग, 2018, पृ. 494-506.

[13] ए. कोनो एट अल., "स्केल पर अनसुपरवाइज्ड क्रॉस-लिंगुअल रिप्रेजेंटेशन लर्निंग," प्रोक. एसीएल, 2020, पृ. 8440-8451.

[14] जी. गोयल और ए. कुलकर्णी, "संस्कृत के लिए वितरणात्मक सिमेंटिक्स," प्रोक. संस्कृत कम्प्यूटेशनल लिंग्विस्टिक्स सिम्पोजियम, 2013, पृ. 191-204.
[15] पी. शार्फ और एम. हाइमन, संस्कृत एन्कोडिंग में भाषाई मुद्दे। प्रोविडेंस, आरआई, यूएसए: ब्राउन यूनिवर्सिटी प्रेस, 2009.
[16] जे. ब्रोंखोर्स्ट, भाषा और वास्तविकता: भारतीय विचार में एक प्रकरण पर। लीडेन, नीदरलैंड्स: ब्रिल, 2011.
[17] वी. गोयल, डी. कुमार, और आर. लेहल, "संस्कृत यौगिक विश्लेषण के लिए एक नियम-आधारित दृष्टिकोण," प्रोक. आईजेसीएनएलपी, 2011, पृ. 1122-1129.
[18] ए. टेरडालकर और ए. भट्टाचार्य, "संस्कृत में प्रश्न-उत्तर के लिए ढांचा," प्रोक. डब्ल्यूएसएसएएनएलपी, 2019, पृ. 24-33.
[19] ओ. हेलविग, "संस्कृत ग्रंथों में वाक्य सीमाओं का पता लगाना," प्रोक. सीओएलआईएनजी, 2016, पृ. 288-297.
[20] पी. कुमार, एस. लाहिड़ी, और ए. बेहेरा, "संस्कृत के लिए डिजिटल संधि विश्लेषक," इंट. जे. कम्प्यूट. एप्लाइड, खंड 41, संख्या. 6, पृ. 15-21, 2012.
[21] टी. बहादुर और पी. झा,"न्यूरल नेटवर्क का उपयोग करके स्वचालित संस्कृत यौगिक विश्लेषण," प्रोक. एलआरईसी, 2020, पृ. 3421-3428 में।
[22] एस. नेहरडिच और ओ. हेलविग, "लैटिन, प्राचीन ग्रीक और वैदिक संस्कृत का सटीक निर्भरता पार्सिंग और टैगिंग," प्रोक. एलटी4एचएएलए कार्यशाला, 2022, पृ. 88-97 में।
[23] ए. भट्टाचार्य, "संस्कृत का मशीन अनुवाद: चुनौतियाँ और दृष्टिकोण," प्रोक. एनएलपीएआई, 2015, पृ. 55-63 में।
[24] आर. अरलिकट्टे एट अल., "seq2seq लर्निंग का उपयोग करके संस्कृत संधि विभाजन," प्रोक. ईएमएनएलपी, 2018, पृ. 4587-4595 में।
[25] जी. ज़िस्क, "डिजिटल मानविकी और संस्कृत पाठ्य छात्रवृत्ति," जे. डिजिटल मानविकी, खंड। 5, नं. 2, पृ. 33-49, 2016.
[26] पी. गोयल और जी. ह्युएट, "संस्कृत कॉर्पस एनोटेशन के लिए एक लीन इंटरफ़ेस का डिज़ाइन और विश्लेषण," जे. लैंग्वेज मॉडलिंग, खंड 4, नं. 2, पृ. 145-182, 2016.
[27] एस. संधान एट अल., "संस्कृत निर्भरता पार्सिंग के लिए एक ट्री रूपान्तरकः मॉडल," प्रोक. ईएमएनएलपी, 2021, पृ. 2569-2579 में.
[28] जे. कौर और आर. शर्मा, "संस्कृत ग्रंथों के लिए शब्द अर्थ अस्पष्टीकरण," प्रोक. आईसीसीएलए, 2019, पृ. 100-107 में.
[29] पी. शुक्ला और ए. कुलकर्णी, "संस्कृत हेरिटेज रीडर: एक उपयोगकर्ता-अनुकूल इंटरफ़ेस," प्रोक. डब्ल्यूएसएसएएनएलपी, 2016, पृ. 45-52.
[30] डी. जुराफस्की और जे.एच. मार्टिन, स्पीच एंड लैंग्वेज प्रोसेसिंग, तीसरा संस्करण। अपर सैडल रिवर, एन.जे., यूएसए: प्रेंटिस हॉल, ड्राफ्ट संस्करण, 2023.

Cite the Article:

मिश्रा,कार्तिकेय, सिंह,डॉ. शालिनी. (2026).आधुनिकसंगणकप्रणाल्याः साहाय्येन संस्कृतभाषायाः व्याकरणिकसंरचनायाः अध्ययनम्. Shiksha Samvad International Open Access Peer-Reviewed & Refereed Journal of Multidisciplinary Research, 3(4) 731-739, ISSN: 2584-0983 (Online), Volume 03, Issue 04, June-2026. **DOI: -** https://doi.org/10.64880/shikshasamvad.v3i4.78

This is an Open Access Journal / article distributed under the terms of the Creative Commons Attribution License CC BY-NC-ND 3.0) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. All rights reserved.

CERTIFICATE

of Publication

This Certificate is proudly presented to

कार्तिकेय मिश्रा[1], डॉ. शालिनी सिंह[2]

For publication of research paper title

आधुनिकसंगणकप्रणाल्याः साहाय्येन संस्कृतभाषायाः व्याकरणिकसंरचनायाः अध्ययनम्

Published in 'Shiksha Samvad' Peer-Reviewed and Refereed Research Journal and E-ISSN: 2584-0983(Online), Volume-03, Issue-04, Month June 2026, Impact Factor-RPRI-3.89/6.69

Dr. Neeraj Yadav
Editor-In-Chief

Dr. Lohans Kumar Kalyani
Executive-chief- Editor

Note: This E-Certificate is valid with published paper and the paper must be available online at: https://shikshasamvad.com/

DOI: - https://doi.org/10.64880/shikshasamvad.v3i4.79