संस्कृत-समास-प्रक्रिया

अनिल-कुमारः[1], विपुल-मित्तलः[2] अम्बा-कुलकर्णी[1]

[1]संस्कृताध्ययनविभागः, हैदराबादविश्वविद्यालयः, हैदराबाद, भारतम्

anil.lalit22@gmail.com, apksh@uohyd.ernet.in

[2]Language Technology Research Centre, IIIT, Hyderabad, India.

vipulmittal@research.iiit.ac.in

सारांशः

संस्कृतभाषायां समासप्रयोगानामाधिक्यं दृश्यते। तत्र समस्तपददर्शनेनैव समासे प्रयुक्ततत्तत्पदस्य सम्बन्धः स्पष्टरूपेण न प्रतीयते। तस्मात् समासार्थप्राप्त्यर्थं प्रथमं समस्तपदे विद्यामानपदेषु परस्परान्वयः ज्ञातव्यः। एवञ्च तद्विग्रहवाक्यमपि प्राप्तव्यम्। प्रकृतसमुद्भतचर्चायां समाससन्दर्भे गणितशास्त्रे उपयुज्यमानसांख्यिकिप्रक्रियासाहाय्येन यन्त्रेण समासच्छेदः(समासे विद्यमानपदानां पृथग्दर्शनम्) समासभेदनिर्धारणञ्च कथमिति पत्रेऽस्मिन् प्रस्तूयते।

1. परिचयः

प्रगतवर्षेषु संस्कृतसङ्गणकभाषाविज्ञानक्षेत्रे बहुविधकार्याणि जातानि। तत्र बहुभिः विद्वद्भिः कृतपरिश्रमैः महता प्रयासेन संस्कृतज्ञानार्थं ([12],[14],[23],[16],[7],[1]) नानाविधसङ्गणकीयोपकरणानि निर्मितानि। संस्कृतपदविश्लेषकः, सन्धिसंश्लेषकः, सन्धिविच्छेदकः इत्याद्युपकरणानि अधिकतया दृश्यन्ते। तत्र कारकाणामुपर्यपि केचन([10],[13],[8]) कार्यमकुर्वन्। समासः इति विषयम् अवलम्ब्य तु समस्तपदविच्छेदपर्यन्तमेव कार्यं दृश्यते।

संस्कृतवाङ्मये समस्तप्रयोगानाम् आधिपत्यं दरीदृश्यते। समासाः प्रयोगदशायां सरलाः, तदर्थं ते असीमिताः। तथा च न सम्भवति सर्वेषां समस्तपदानां शब्दकोशे योजनं तत्प्रयोगाधिक्यात् तद्गानार्थकत्वाच्च। समाससंरचनायां सन्धिः[1] नित्यः। समासपदे सन्धिविच्छेदेनैव तत्पदार्थः नोपलभ्यते, तस्मिन्विद्यमानपदानां परस्परान्वयाभावात्। समस्तपदार्थावबोधनाय सर्वप्रथमं समस्तपदे विद्यमानपदानां परस्परान्वयः ज्ञातव्यः, तस्मात् विग्रहवाक्यं[2] स्वयमेवोपलभ्यते। 'कोन्टेक्स्ट फ्री ग्रामर् ' (Context free grammar) इति आधारीकृत्य समासेषु कथं विभक्तिरित्यादिविशेषाः समीचीनरीत्या अनुबन्धितव्याः इति गिल्लन् (Gillon) [6] महोदयः स्पष्टयति। सः समासपदेषु विभक्तीनां, तत्रविद्यमानविशेष्यम् , समासस्थघटकानाञ्च

---

1 संहितैकपदे नित्या नित्याधातूपसर्गयोः। नित्या समासे वाक्ये सा विवक्षामपेक्षते।।

2 समासस्य यः अर्थः तमेव विग्रहवाक्यम् इति उच्यते।

अनुबन्धयति। स च सूचयति यत् कुत्रचित् समासघटकैः यथा 'न' इत्यनेन समासघटकेनैव तत्तत्समासभेदस्य परिचयनम् कथम् इति। इदं हि दर्शयति भगवान् पाणिनेः सूक्ष्मबुद्दिमत्त्वम्। तेन एतादृशांशा अपि सूत्रिताः। यन्त्रस्य सन्दर्भज्ञानं न विद्यते इति हेतोः एतत्सर्वस्यापि यान्त्रिकीकरणमधुनापि दूरं स्थीयते। परन्तु सन्दर्भं विनापि सांख्यिकिप्रक्रियामुपयुज्य तत् साधयितुं शक्यते इत्यङ्गीकुर्वन्ति सङ्गणक-भाषावैज्ञानिकाः। सांख्यिक्युपकरणानि मानवीयानुबन्धिकृतोदाहरणानि अधिकृत्य (manaully tagged data) यन्त्रं शिक्षयन्ति। अनन्तरमेतदेव शिक्षणं भाषायाः अध्ययनार्थमुपयुज्यते।

एवमेव समासस्य मानवीयानुबन्धिकृतोदाहरणानि स्वीकृत्य सामान्यसांख्यिकिप्रक्रियामुपयुज्य कथं समासपदच्छेदः (समासे विद्यमानपदानां पृथग्दर्शनम्) समासभेदनिर्धारणञ्च यन्त्रेण कर्तुं शक्यत इत्येतत् विषयं स्वीकृत्य चर्चा अग्रे भविष्यति।

2. संस्कृतसमासः

कस्तावत् समास इति पदार्थः इति विचिकित्सायां 'समसनमिति समासः' इति प्रोक्तम्। बहुपदानां एकपदीकरणं, समस्तपदे यः अर्थः सैव विग्रहवाक्ये भवितव्यम् इति समासपदेनोच्यते। समासप्रक्रियायां विभक्तिलोपादिकार्याणि विधीयन्ते तेषां विवरणञ्च अधः विवृण्यते।

1) समासः सर्वदा ऐकपदं[3] स्यात् ।

2) समासः ऐकविभक्तिकं[4] स्यात्। अलुक्समासे तु अपवादरूपेण पूर्वपदे विभक्तिलोपः न भवति यथा युधिष्ठिरः इत्यादौ युधि इत्यत्र न विभक्तिलोपः।

3) समासः ऐकस्वर्यं[5] स्यात्।

4) समासे विद्यमानपदानां क्रमपरिवर्तनं न स्यात्।[6]

5) समासे व्यवधानं न स्यात्।[7]

6) द्वन्द्वसमासः बहुपदबहुव्रीहिसमासश्च विहाय अन्यसमासेषु द्वयोर्पदयोरेव समासः विधीयते।

7) नित्यः सन्धिः समासेषु।

8) समासे विद्यमानपदानां स्वलिङ्गं स्ववचनञ्च विहाय पृथग्लिङ्गं पृथग्वचनञ्च भवितुम् अर्हति। यथा - पाणिपादम् , पाचिकाभार्यः इति।

संस्कृतवाङ्मये द्विपदात्मक-त्रिपदात्मकसमासप्रयोगाः बहुधा दृश्यन्ते। अपि च बहुषु स्थलेषु बहुपदात्मकसमस्तप्रयोगाः विरलाः न, संस्कृतभाषायाम्। अधः प्रस्तूयन्ते कानिचन् उदाहरणानि ।

3 ऐकपद्यं ऐकस्वर्यं च समासवत् भवति [काशिका – 2.1.46], समर्थः पदविधिः [महाभाष्यम् – 2.1.1]

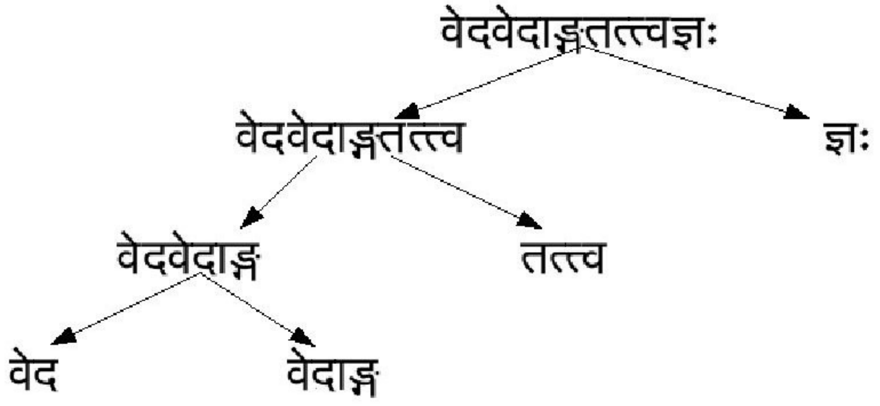
4 ऐकपद्यं ऐकस्वर्यं च समासवत् भवति [काशिका – 2.1.46], समर्थः पदविधिः [महाभाष्यम् – 2.1.1]

5 समर्थः पदविधिः [महाभाष्यम् – 2.1.1] इति सूत्रे वर्णितम्।

6 समर्थः पदविधिः [महाभाष्यम् – 2.1.1] इति सूत्रे वर्णितम्।

7 समर्थः पदविधिः [महाभाष्यम् – 2.1.1] इति सूत्रे वर्णितम्।

1) वेदवेदाङ्गतत्त्वज्ञः[8] = वेद-वेदाङ्ग-तत्त्व-ज्ञः



चित्रम् 1 - वेदवेदाङ्गतत्त्वज्ञः इत्यस्य पदान्वयप्रदर्शनम्

बहुपदसमासप्रयोगे द्विपदात्मकसमासप्रयोगानां समावेशः प्रकृतचित्रे -1 द्रष्टुं शक्यते।

2) प्रवरमुकुटमणिमरीचिमञ्जरीचयचर्चितचरणयुगलम्[9] = प्रवर-मुकुट-मणि-मरीचि-मञ्जरी-चय-चर्चित-चरण-युगलम् ।

3) जलादिव्यापकपृथिवीत्वाभावप्रतियोगिपृथिवीत्ववती[10] = जल-आदि-व्यापक-पृथिवीत्व-अभाव-प्रतियोगि-पृथिवीत्ववती ।

आर्थिकदृष्ट्या भगवता पाणिनिना समासः चतुर्विधः प्रोक्तः। यथा -

1) तत्पुरुषः (प्रायेण उत्तपदार्थप्रधानः)

2) बहुव्रीहिः (प्रायेण अन्यपदार्थप्रधानः)

3) द्वन्द्वः (प्रायेण उभयपदार्थप्रधानः)

4) अव्ययीभावः (प्रायेण पूर्वपदार्थप्रधानः)

एते भेदाः सर्वेषां समस्तपदानां विग्रहवाक्यं निर्मातुं न पर्याप्ताः । यथा 'वृक्षमूलम् ' इत्यस्य विग्रहवाक्यं 'वृक्षस्य मूलम् ' इति एवञ्च 'गृहगतः' इत्यस्य विग्रहवाक्यं 'गृहं गतः' इति च स्यात् । एतदुभयमपि तत्पुरुषसमासस्योदाहरणम्। परन्तु उभयोरपि विग्रहवाक्यं दृष्ट्वा ज्ञायते यत् षष्ठीतत्पुरुषः, द्वितीयातत्पुरुषः इत्यादिरूपेण उपरितनप्रोक्तप्रतिसमासस्य अन्यभेदाः अपि विद्यन्ते। पृथक्-पृथक् समासानां विग्रहवाक्यानि दृष्ट्वा अस्माभिः 55 समासस्य अन्यभेदाः प्राप्ताः। समासस्य अन्यभेदान् द्रष्टुं कृपया परिशिष्टः दृश्यताम्। परिशिष्टः-अ मध्ये समासस्य इतरभेदानां विवरणं विद्यते। परिशिष्टः-ब मध्ये विग्रहवाक्यनियमाः विद्यन्ते।

8 रामायणम् – 1.1.14

9 पञ्चतन्त्रे कथामुखे प्रोक्तम् ।

10 केवलव्यतिरेकी–प्रकरणम् : मणिकणः [22]

३. समासप्रक्रिया

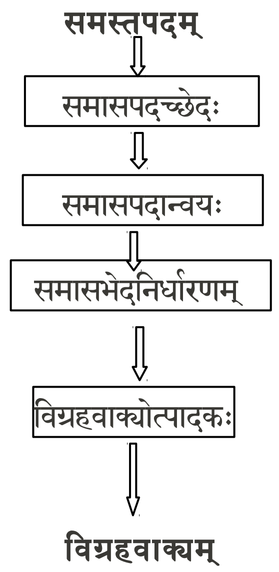
समासप्रक्रिया चतुर्सोपानेषु विभक्ता। तानि सोपानानि यथा -

1) समासपदच्छेदः

2) समासपदान्वयः

3) समासभेदनिर्धारणम्

4) विग्रहवाक्यञ्चेति।



3.1 समासपदच्छेदः

सोपानेऽस्मिन् समासे विद्यमानपदानां पृथक्करणं विधीयते। यथा -

सुमित्रानन्दवर्धनः

इत्यस्य पृथक्करणं

सुमित्रा-आनन्द-वर्धनः

इति।

3.2 समासपदान्वयः

समासपदच्छेदानन्तरं ततः प्राप्तपदेषु परस्परान्वयः विधीयते । तथा च द्वयोर्द्वयोः पदयोः समूहीकरणञ्च सोपानेऽस्मिन् विधीयते। यथा -

सुमित्रा-आनन्द-वर्धनः

इत्यस्य समूहीकरणं

<सुमित्रा-<आनन्द-वर्धनः>>

इति।

3.3 समासभेदनिर्धारणम्

समासपदान्वयानन्तरं ततः प्राप्तसमूहानां समासभेदनिर्धारणं विधीयते। यथा -

<सुमित्रा-<आनन्द-वर्धनः>>

इत्यस्य समासनामनिर्धारणं

<सुमित्रा-<आनन्द-वर्धनः>T6>T6

इति। अत्र 'T6' इति षष्ठीतत्पुरुषसमासः इति सूचयति। एतत्सोपानाय आर्थिकपक्षस्यापि अपेक्षा विद्यते।

3.4 विग्रहवाक्यम्

समासभेदनिर्धारणानन्तरं ततः एकैकं समूहम् अवलम्ब्य विग्रहवाक्यम् उत्पद्यते एतस्मिन्सोपाने। यथा -

आनन्दस्य वर्धनः = **आनन्दवर्धनः**

सुमित्रायाः आनन्दवर्धनः = **सुमित्रानन्दवर्धनः**

इति।

4. समासपदच्छेदः

'एकं समस्तपदं स्वीकृत्य तत्रस्थपदानां पृथग्दर्शनमेव समासपदच्छेदः' इत्युपकरणस्य कार्यम्। समासेषु सन्धिः नित्यः। एवञ्च समासच्छेदार्थं सन्धिविच्छेदः प्रथमं सोपानं भवति। सन्धिविच्छेदार्थं सन्धिनियमाः अपेक्षन्ते। सन्धिनियमेषु अस्माभिः पक्षत्रयं स्वीकृतम् । तत्र प्रथमपक्षे सन्धिवर्णानां युगलं स्यात् यथा - रामश्चिनोति इत्यत्र '**श्च**' इति वर्णद्वययुगलम्। द्वितीयपक्षे प्रथमशब्दस्यान्तिमाक्षरं स्यात् यथा - रामश्चिनोति इत्यत्र प्रथमं पदं '**रामस्**' इत्यस्यान्तिमाक्षरं '**स्**' इति। तृतीयपक्षे द्वितीयशब्दस्यादिमक्षरं स्यात् यथा - रामश्चिनोति इत्यत्र द्वितीयं पदं '**चिनोति**' इत्यस्यादिममक्षरं '**च**' इति । सर्वेऽपि सन्धिनियमाः 'प्रथमपक्षम् , द्वितीयपक्षम् , तृतीयपक्षम् ' इत्यादिरूपेण समानान्तरेण व्यवस्थिताः। अधुना सन्धिविच्छेदनार्थं प्रथमपक्षमाश्रित्य '**द्वितीयपक्षम् -+ तृतीयपक्षम्**' इति रूपेण विभज्यते यथा - '**श्च**' इत्यस्य '**स्-+च**' इति रूपेण विभाजनम् । एवम् अनुसरणेन तत्र बहुविकल्पानि सन्धिविच्छेदरूपाणि अपेक्षितानपेक्षितरूपाणि च व्युत्पाद्यन्ते। तस्मात् सांख्यिकिप्रक्रियासाहाय्येन अपेक्षितसन्धिविच्छेदितरूपं स्वीक्रियते। एतादृशीं प्रक्रियां भाषावैज्ञानिकाः 'ओप्टीमलिटि थ्योरी (Optimality Theory)' इति नाम्ना व्यवह्रियन्ते। एतत्प्रक्रियायां सोपानत्रयम् आश्रित्य कार्यं विधीयते। सोपानत्रयं नाम 'व्युत्पादनम् - अवरोधनम् - मूल्याङ्कनम् (GENerate – CONstrain – EVALuate)' इति। एतत्सोपानत्रयम्

अस्माभिरपि अनुसर्यते समस्तपदच्छेदनार्थम् । वस्तुतः एतस्याः प्रक्रियायाः उपयोगः व्युत्पादनार्थम् (Generation) उपयुज्यते। केचन([4],[3]) एतस्याः प्रक्रियायाः उपयोगः विपरीतदृष्ट्या अपि अकुर्वन् अपि च सफलताम् अप्राप्नुवन्।

अष्टाध्यायीग्रन्थस्थसन्धिविधायकसूत्राणाम् उपयोगः अनपेक्षितरूपान् अवरोद्धुं अर्थात् अवरोधनम् (CONstrain) इति सोपानरूपेण क्रियते। तथा च तेषामेव सन्धिविधाकयकसूत्राणाम् उपयोगः सन्धिविच्छेदनार्थम् अपि करणीयश्चेत् तदेकं जटिलं कार्यं स्यात्। यतोहि सन्धिविधायकसूत्राणि तु केवलं सन्धिः कुत्र आचरणीयः इत्येव सूचयन्ति, न तु सन्धियुक्तपदं कथं विच्छेदनीयम् इति। 'ओप्टीमलिटि थ्योरि (Optimality theory)' इत्यस्याः प्रक्रियायाः सोपानानि अधः विस्तृतरूपेण विवृण्यन्ते।

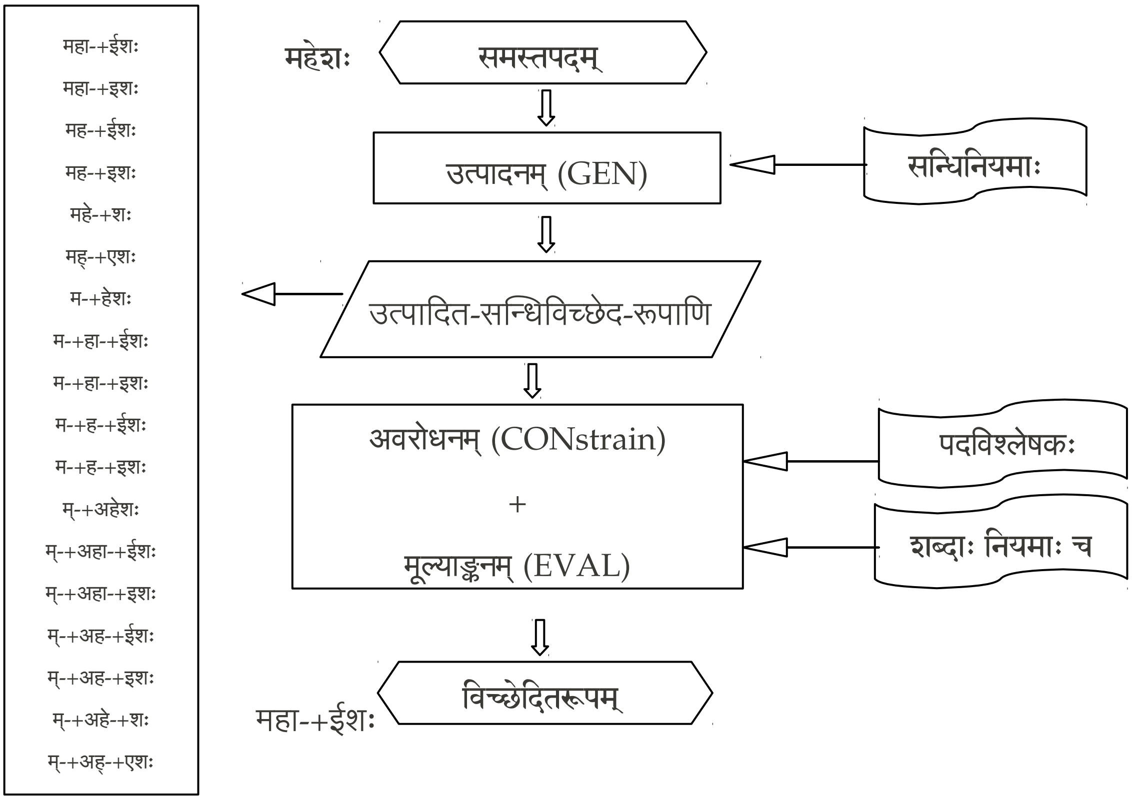
4.1 स्कोरिंग मैट्रिक्स (Scoring Matrix)

सन्धिसहित-सन्धिरहितोदाहरणानि अस्माभिः भारतसर्वकारस्य 'संस्कृत-समूह-परियोजना (Sanskrit consortium project)' इत्यस्मात् सङ्गृहीतानि। तानि उदाहरणानि विभिन्नग्रन्थेभ्यः स्वीकृतानि। ग्रन्थेषु बालबोधिका, बालकथा, नाटकानि, पुराणानि, आयुर्वेदग्रन्थाः प्रमुखाः सन्ति। एतेभ्यः ग्रन्थेभ्यः शतसहस्रशब्दाः स्वीकृताः। तेषु पञ्चविंशतिसहस्रात्मकानि सन्धियुक्तोदाहरणानां सन्धिविच्छेदनं कारयित्वा समानन्तररूपेण व्यवस्थितानि। एतेषाम् उदाहरणानां सन्धिनियमानाम् आवृत्तिसङ्ख्यां प्राप्तुम् उपयोगः कृतः। यथा सर्वणसन्धिविधायकोदहारणानि कति, कियद्वारञ्च आगतानि इत्यादिरूपेण सन्धिनियमाः व्यवस्थिताः । यदि अपेक्षितसन्धिनियमः न प्राप्यते तर्हि तस्य सामान्यरूपेण 1 इति आवृत्तिसङ्ख्यां विचिन्त्य कार्यं सम्पादयामः इति सिद्धान्तः।

4.2 समासपदच्छेदप्रक्रिया सोपानम् च

यथा पूर्वमेव विचारितं यत् समासपदच्छेदार्थम् अस्माभिः 'ओप्टीमलिटि थ्योरि' (Optimality Theory) इत्यस्याः प्रक्रियायाः उपयोगः क्रियते। तथा च तत्र सोपानत्रयम् '**उत्पादनम् - अवरोधनम् - मूल्याङ्कनम्** ' इत्येतत् आधारीकृत्य कार्यं विधीयते। उत्पादनम् इत्यस्मिन् सोपाने यावच्छक्यं तावत् अपेक्षित-अनपेक्षित-समासविच्छेदरूपाणाम् उत्पादनं भवति। तदनन्तरं 'पदविश्लेषकः' इत्युपकरणस्य साहाय्येन अनपेक्षितसमासविच्छेदरूपाणि निष्कास्यन्ते।

अवरोधनम् -मूल्याङ्कनम् च इत्यनयोः सोपानयोः अनपेक्षितसमासविच्छेदितरूपाणां निष्कासनम् अवरोधनञ्च भवति। तत्र अवरोधनं निष्कासनञ्च द्विविधा क्रियते। तद्यथा - 1) पदविश्लेषकः इत्युपकरणम् उपयुज्य समस्तपूर्वपद-उत्तरपदयोः साधुत्वं स्वीक्रियते। यदि कस्यचित् पदस्य असाधुत्वं प्राप्यते चेत् तद् समस्तपदं सम्यक् न विच्छेदितम् इति विचिन्त्य तस्य तत्रैव त्यागः करणीयः इति । 2) सर्वेषाम् समासविच्छेदितरूपाणाम् यद् अन्तिमं पदं स्यात् तेनैव समासविभक्त्यादिविचारः इति। उदाहरणार्थं अधः प्रदत्तं चित्रं दृश्यताम्।



समासपदविच्छेदप्रक्रियायाः सोपानानि

1. सन्धिनियमान् उपयुज्य प्रदत्तसमस्तपदस्य सर्वं सम्भावितं विच्छेदनं प्राप्तव्यम्।
2. पदविश्लेषकः इत्युपकरणम् उपयुज्य समस्तपूर्वपद-उत्तरपदयोः साधुत्वं प्राप्तव्यम्।
3. पूर्वस्मिन् प्रघट्टे यथा वर्णितं तथा प्रति समस्तपदस्य व्यवहारदृष्ट्या श्रेणिप्रदानम् ।
4. श्रेणिप्रदानानन्तरम् सर्वाणि समस्तरूपाणि क्रमशः स्थापनीयानि। तत्र प्रथमश्रेण्यां यत् समस्तं तस्यैव प्राधान्यम् इति।

परिणामः

पदविश्लेषकः[11] इत्युपकरणं चतुर्दशकोटिशब्दानां विश्लेषणं कर्तुं समर्थः। आहत्य 2,650 सन्धिनियमाः सङ्गृहीताः। तद्परीक्षणार्थं 8,260[12] उदाहरणानि समानन्तररूपेण स्वीकृतानि। परीक्षणानन्तरम् अस्माभिः यः परिणामः प्राप्तः। सः अधः प्रस्तूयते।

-प्रथमश्रेण्यां यानि समस्तपदविच्छेदितरूपाणि, तानि 92.5% समीचीनानि सन्ति। श्रेण्याम् आरम्भिकत्रिषु आगतानि यानि रूपाणि तानि 91.1% समीचीनानि सन्ति।

-सामान्यतया 15 अक्षरात्मकशब्दानां विच्छेदनार्थं यन्त्रं 0.04 क्षणात्मककालं स्वीकरोति।[13]

अधः प्रदत्ततालिकायां श्रेणिप्रदानविवरणमित्यादि दर्श्यते।

11 http://sanskrit.uohyd.ernet.in/~anusaaraka/sanskrit/samsaadhanii/morph/index.html

12 परिक्षणार्थम् उदाहरणानि महाभारतम् इति ग्रन्थात् स्वीकृतानि।

13 अस्माभिः 2.93GHz Core 2 Duo processor and 2GB RAM इति यन्त्रे परिक्षितम्।

श्रेणी	%
1	92.46 35
2	5.05 92
3	1.62 35
4	0.29 78
5	0.19 36
>5	0.37 23
श्रेण्यानुसारेण प्रतिशद्दर्शनम्	

5. समासपदान्वयः

समासपदच्छेदः पदविश्लेषकस्य सन्धिच्छेदकस्य च साहाय्येन एकस्य समस्तपदस्य एकं वा एकाधिकं वा छिन्नरूपं जनयति। तद्छिन्नरूपं स्वीकृत्य समासपदान्वयः द्वयोर्द्वयोः पदयोः समूहीकरणं करोति। एतत्समूहीकरणेन ज्ञायते यत् कियद्पर्यन्तं छिन्नरूपाणां समूहीकरणं कर्तुं शक्यत इति। उदाहरणार्थं त्रिपदात्मक-छिन्न-समस्तपदं वर्तते यथा पद 1-पद 2-पद 3 इति । सामान्यसमस्तपदे पदद्वयोरेव समावेशः दृश्यते । तथा चेत् पद 1-पद 2-पद 3 इत्यत्र द्विविधं समूहविभजनं कर्तुं शक्यते। यथा - <पद 1<पद 2-पद 3>> अथवा <<पद 1-पद 3> पद 3> इति। प्रसङ्गदृष्ट्या तु तयोः एक एव समीचिनः समूहः । स्पष्टार्थं प्रदत्ते उदाहरणे दृश्येताम्।

1) <बद्ध-<जाम्बूनद-स्रजः>>

2) <<तपस् -स्वाध्याय>-निरतम् >

त्रिपदात्मकसमस्तपदे तु केवलं द्विविधं समूहीकरणं सम्भवति इति दृष्टम्। परन्तु समस्तपदे यदा पदानां संख्या वर्धते तदा तु समूहानां संख्यापि एधिष्यते। यथा - यदि चतुष्पदात्मकः समासः स्यात् तर्हि तस्य पञ्चप्रकारेण समूहीकरणं सम्भवति। यदि पञ्चपदात्मकः समासः तर्हि चतुर्दशप्रकारेण समूहीकरणं सम्भवति। यदि षष्ठपदात्मकः समासः तर्हि तस्य द्वाचत्वारिंशत्प्रकारेण तत्समूहीकरणं सम्भवति इति। गणितशास्त्रे एतदर्थं साधारणरूपेण उच्यते यदि $n+1$ पदानि स्युः तदा तु 'केटालन् नम्बर' (Catalan number) इति सिद्धान्तम् उपयुज्य समूहीकरणं कर्तुं शक्यते इति। सः सिद्धान्तः एवं वर्तते -

$$Cn = (2n)!/(n+1)!n!$$

इति। सन्दर्भदृष्ट्या समीचीनसमूहस्य चयनमेव समासपदान्वयस्य कार्यम् इति । वस्तुतः एतादृशं कार्यं यान्त्रिकदृष्ट्या अतिमहत्त्वपूर्णं वर्तते। एतस्यां दिशि अपि कार्यं करणीयम्।

6. समासभेदनिर्धारणम्

समासपदान्वयः एकवारं लभ्यते चेत् तदनन्तरम् एकैकं समूहं स्वीकृत्य तत्र कीदृशः समासः भवितुमर्हति इति निर्धारणमेव 'समासभेदनिर्धारकः' इत्युपकरणस्य कार्यम् इति। पूर्वोक्तरीत्या अस्माभि 55 समासानाम् द्वितीयातत्पुरुषेत्यादिरूपेण भेदाः सङ्गृहीताः। तत्र प्रतिभेदाय अनुबन्धनाम अपि प्रदत्तम् यथा T6 (षष्ठीतत्पुरुषः), Bs2 (द्वितीयाबहुव्रीहिः) इति।

एतत्कार्यसम्पादनाय अस्माभिः मानवीयानुबन्धितोदाहरणानि स्वीकृतानि। तत्र आहत्य 150 सहस्रपदानि वर्तन्ते। तेषु 12,796 समस्तपदानि[14] विद्यन्ते। एतानि उदाहरणानि परिशिष्टः-1 मध्ये प्रदत्तानुबन्धसूचिकानुसारेण अनुबन्धितानि। तत्र विद्यमानानि उदाहरणानि सन्दर्भदृष्ट्या अनुबन्धितानि। एतानि उदाहरणानि यन्त्रस्य प्रशिक्षणार्थं स्वीकृतानि। अन्यानि च 400 उदाहरणानि परीक्षणार्थं स्वीकृतानि।

मानवीयानुबन्धितोदाहरणानां विषये केचन अंशाः -

1) 10% उदाहरणानां पुनरावृत्तिः अपि वर्तते।

2) 12,976 समस्तपदेषु 2,630 विभिन्नानि पूर्वपदानि विद्यन्ते तथा च 7,106 विभिन्नानि उत्तरपदानि विद्यन्ते।

3) तालिका-2 मध्ये समासस्य अन्यभेदानां आवृत्तिसंख्याविषयकसूचनाः विद्यन्ते। तालिका-3 मध्ये आर्थिकदृष्ट्या समासभेदानां आवृत्तिसंख्याविषयकसूचनाः विद्यन्त इति।

अस्माभिः आवृत्तिसंख्यां प्राप्तुं अधः प्रदत्तसांख्यिकिनियमः प्रयुज्यते। तद्यथा -

P(T/पद 1-पद 2) = पद 1-पद 2 इति समासस्य, 'T' इति समासभेदः सम्भवितुम् अर्हति इति।

समासभेदः	शब्दाः (%)
T6	28.35
Bs6	12.45
K1	9.63
Tn	8.56
Di	7.23
U	5.73

तालिका-2

समासभेदः	शब्दाः (%)
T	52.43
B	18.96
K	12.04
D	8.84
U	5.73

तालिका-3

तर्हि

14 सामान्यसमस्तानि उदाहरणानि। यत्र द्वयोः द्वयोः पदयोः एव समासः।

$$P\ (T/\text{पद}\,1\text{-पद}\,2) = P(T/\text{पद}\,1)*P(T/\text{पद}\,2)$$

यत्र

$$P(T/\text{पद}\,1) = P(T*\text{पद}_i)/P(\text{पद}_i),\ i=1.2$$

यत्र तु उपरितननियमैः किमपि न प्राप्यते तत्र अधः प्रदत्तनियमः प्रयुज्यते।

$$P(T.\text{पद}\,i) = 1/F,$$

$P(\text{पद}_i) = 1/F$ अत्र F इत्यक्तौ समासमधिकृत्य मानवीयानुबन्धितकृतशब्दानां सङ्ख्या।

6.1 मूल्याङ्कनम्

मूल्याङ्कनार्थम् अस्माभिः 400 समासोदाहरणानि स्वीकृतानि। तत्रापि यत्र द्वयोः पदयोरेव समासः तादृशोदाहरणानि एव स्वीकृत्य सन्दर्भं दृष्ट्वा अनुबन्धितानि। समासनामनिर्धारक इत्युपकरणन्तु सन्दर्भं न जानाति। तदर्थं तद् एकाधिकम् उत्तरं प्रदर्शयति। किम् उत्तरम् आदौ, किमनन्तरं च भवितव्यम् इति तादृशं क्रममपि निर्माति। सामान्यतया मूल्याङ्कने प्रदत्तोत्तराणां शुद्धता सीमा च दृष्टव्या। परन्तु समासनामनिर्धारकस्तु प्रदत्तसमासपदस्य आवृत्तिसंख्यां दृष्ट्वा समासभेदं निर्धारयति। अत्र प्रदत्तक्रमः मुख्यः भवति। यतोहि क्रमं दृष्ट्वा एवास्माभिः तस्य शुद्धता इत्यादिनां विषये चिन्तयितुं शक्यते। अधः सामान्यरूपेण अस्माभिः मूल्याङ्कनं तालिकायां प्रदर्श्यते।

	55 समासभेदानां मूल्याङ्कनम्		8 समासभेदानां मूल्याङ्कनम्	
क्रमः	**शब्दानां सङ्ख्या**	**शब्दाः (%)**	**शब्दानां सङ्ख्या**	**शब्दाः (%)**
1	252	63	291	72.7
2	44	10.9	53	13.2
3	29	7.2	38	9.5

तालिका-4 समासनामनिर्धारकस्य मूल्याङ्कनसूचनाः

यदि वयं प्रथमक्रमस्य विद्यमानमूल्याङ्कनमेव स्वीकुर्मः चेत् तदा समासभेदनिर्धारकस्य 55 समासभेदानां शुद्धता 63% वर्तते, यदि केवलं 8 समासभेदानां स्वीकुर्मश्चेत् तदा समासभेदनिर्धारकस्य शुद्धता 72.7%. परन्तु यदि क्रमत्रयमपि स्वीकुर्मश्चेत तदानीं तस्य शुद्धता 95% भवति।

एतस्य कार्यदक्षतां सन्दर्भनियमान् संयोज्य इतोपि एधयितुं शक्यते। अष्टाध्यायां सामान्यतया 200 समासविधायकसूत्राणि विद्यन्ते । यथा - अन्नेन व्यञ्जनम् [2.1.34] इत्यनेन सूत्रेण व्यञ्जनवाचितृतीयान्तसुबन्तस्य अन्नवाचिसुबन्तेन सह विकल्पेन तत्पुरुषसमासः विधीयते । संस्कृत-समास-प्रक्रिया इत्यस्य अन्तर्जालसंस्करणमपि विद्यते। तद्द्रष्टुं कृपया अधः प्रदत्तसङ्केतः प्रयुज्यताम्।

http://sanskrit.uohyd.ernet.in/samAsa/frame.html

References

1. Bharati, A., Kulkarni, A.P., and Sheeba, V.: *Building a wide coverage Sanskrit Morphological Analyser: A practical approach.* In: The First National Symposium on Modelling and Shallow Parsing of Indian Languages, IIT-Bombay. (2006)

2. Bhat, G.M.: *Samāsaḥ* . Samskrita Bharati, Bangalore, Karnataka, 2006.

3. Fortes, F.C.L., and Roxas, R.E.O.: *Optimality Theory in Morphological Analysis* In: National Natural Language Processing Research Symposium, January 2004.

4. Fosler, J.E.: *On Reversing the Generation Process in Optimality Theory.* In : Proceedings of the Association for Computational Linguistics. (1996) .

5. Gillon, B.S.: *Exocentric Compounds in Classical Sanskrit.* In: Proceeding of the First International Symposium on Sanskrit Computational Linguistics(SCLS-2007), Paris, France, 2007.

6. Gillon, B.S.: *Tagging Classical Sanskrit Compounds*. In: Sanskrit Computational Linguistics 3, pages 98-105, Springer-Verlag LNAI 5406. (2009)

7. Hellwig, O.: *Sanskrit Tagger: A Stochastic Lexical and POS Tagger for Sanskrit*. In: Sanskrit Computational Linguistics 1 & 2, pages 266-277, Springer-Verlag LNAI 5402. (2007)

8. Hellwig, O.: *Extracting Dependency Trees from Sanskrit Texts*. In: Sanskrit Computational Linguistics 3, pages 106-115, Springer Verlag LNAI 5406. (2009)

9. Hoeks, J.C.J., and Hendriks, P.: *Optimality Theory and Human Sentence Processing: The Case of Coordination.* In: Proceedings of the 27th Annual Meeting of the Cognitive Science Society, Erlbaum, Mahwah, NJ, pp. 959–964. (2005)

10. Huet, G.: *Shallow syntax analysis in Sanskrit guided by semantic nets constraints*.
In: Proceedings of International Workshop on Research Issues in Digital Libraries, Kolkata. (2006)

11. Huet, G.: Lexicon-directed Segmentation and Tagging of Sanskrit. XIIth World Sanskrit Conference, Helsinki, Finland, Aug. 2003. In Themes and Tasks in Old and Middle Indo-Aryan Linguistics, Eds. Bertil Tikkanen and Heinrich Hettrich. Motilal Banarsidass, Delhi, 2006, pp. 307-325.

12. Huet, G.: *Formal structure of Sanskrit text: Requirements analysis for a Mechanical Sanskrit Processor*. In: Sanskrit Computational Linguistics 1 & 2, pages 162-199, Springer-Verlag LNAI 5402. (2007)

13. Huet, G.: *Sanskrit Segmentation.* In: South Asian Languages Analysis Roundtable XXVIII, Denton, Ohio. (October 2009)

14. Jha, G.N. and Mishra, S.K.: *Semantic Processing in Pāṇini's kāraka System*. In: Sanskrit Computational Linguistics 1 & 2, pages 239-252, Springer-Verlag LNAI 5402. (2007)

15. Kulkarni, A.P., Shukla, D.: *Sanskrit Morphological Analyser*: Some Issues. In: To appear in Bh.K

Festschrift volume by LSI. (2009)

16. Kulkarni, A.P., Kumar, A., Sheeba, V.: *Sanskrit compound paraphrase generator*. In: Proceedings of ICON-2009: 7th International Conference on Natural Language Processing, Macmillan Publishers, India.

17. Mahavira: *Pāṇini as Grammarian* (With special reference to compound formation). Bharatiya Vidya Prakashan [Delhi - Varanasi], India, June 1978.

18. Mittal, V.: *Automatic Sanskrit Segmentizer using Finite State Transducers*. In: Proceeding of Association for Computational Linguistics - Student Research Workshop. (2010)

19. Murty, M.S.: *Sanskrit Compounds-A Philosophical Study.* Chowkhamba Sanskrit Series Office, Varanasi(India), 1974.

20. Pandit Ishvarachandra: Aṣṭādhyāyī Chaukhamba Sanskrit Pratisthan, Delhi, 2004.

21. Prince, A., Smolensky, P.: *Optimality Theory: Constraint Interaction in Generative Grammar.* In: RuCCS Technical Report 2 at Center for Cognitive Science, Rutgers University, Piscataway. (1993)

22. Sarma, E.R.S.: *Maṇikaṇa : A Navya-Nyāya Manual.* The Adyar Library and research Centre, Madras. (1960)

23. Scharf, P.M.: *Levels in Pāṇini's Aṣṭādhyāyī* In: Sanskrit Computational Linguistics 3, pages 66-77, Springer-Verlag LNAI 5406. (2009)

24. Yuret, D., Bicici, E.: *Modeling Morphologically Rich Languages Using Split Words and Unstructured Dependencies.* In: ACL-IJCNLP, Singapore. (2009)

परिशिष्टः-अ : संस्कृतसमासभेदाः

अव्ययीभावः		कर्मधारयः	
समासभेदाः	**अनुबन्धनाम**	**समासभेदाः**	**अनुबन्धनाम**
अव्यय-पूर्वपदः	A1	विशेषण-पूर्वपद-कर्मधारयः	K1
अव्यय-उत्तरपदः	A2	विशेषण-उत्तरपद-कर्मधारयः	K2
तिष्ठद्गुप्रभृतिः	A3	विशेषण-उभयपद-कर्मधारयः	K3
संख्यापूर्वद-नद्युत्तरपदः	A4	उपमान-पूर्वपद-कर्मधारयः	K4
नद्युत्तरपद-अन्यपदार्थसंज्ञायाम्	A5	उपमान-उत्तरपद-कर्मधारयः	K5
संख्यापूर्वपद-वंश्योत्तरपदः	A6	अवधारणापूर्वपद-कर्मधारयः	K6
पारे-मध्ये-पूर्वपदषष्ठ्युत्तरपदः	A7	सम्भावनापूर्वपद-कर्मधारयः	K7
		मध्यमपदलोपिकर्मधारयः	Km
तत्पुरुषः		**बहुव्रीहिः**	
समासभेदाः	**अनुबन्धनाम**	**समासभेदाः**	**अनुबन्धनाम**
प्रथमातत्पुरुषः	T1	द्वितीयार्थबहुव्रीहिः (समानाधिकरणः)	Bs2
द्वितीयातत्पुरुषः	T2	तृतीयार्थबहुव्रीहिः (समानाधिकरणः)	Bs3
तृतीयातत्पुरुषः	T3	चतुर्थ्यर्थबहुव्रीहिः (समानाधिकरणः)	Bs4
चतुर्थीतत्पुरुषः	T4	पञ्चम्यर्थबहुव्रीहिः (समानाधिकरणः)	Bs5
पञ्चमीतत्पुरुषः	T5	षष्ठ्यर्थबहुव्रीहिः (समानाधिकरणः)	Bs6
षष्ठीतत्पुरुषः	T6	सप्तम्यर्थबहुव्रीहिः (समानाधिकरणः)	Bs7
सप्तमीतत्पुरुषः	T7	दिग्वाचक-बहुव्रीहिः (समानाधिकरणः)	Bsd
नञ्तत्पुरुषः	Tn	प्रहरणविषयक-बहुव्रीहिः (समानाधिकरणः)	Bsp
समाहारद्विगुः	Tds	ग्रहणविषयक-बहुव्रीहिः (समानाधिकरणः)	Bsg
तद्धितार्थद्विगुः	Tdt	अस्त्यर्थ-मध्यमपदलोपि(नञ्)-बहुव्रीहिः	Bsmn
उत्तरपदद्विगुः	Tdu	प्रादि-बहुव्रीहिः	Bvp
गतिसमासः	Tg	संख्योभयपद-बहुव्रीहिः (समानाधिकरणः)	Bss

कुसमासः	Tk	उपमानपूर्वपद-बहुव्रीहिः (समानाधिकरणः)	Bsu
प्रादिसमासः	Tp	व्यधिकरण-बहुव्रीहिः	Bv
मयूरव्यंस्कादिसमासः	Tm	सङ्ख्योत्तरपद-व्यधिकरण-बहुव्रीहिः	Bvs
तत्पुरुषः बहुपदः	Tb	सहपूर्वपद-व्यधिकरण-बहुव्रीहिः	BvS
तत्पुरुषः उपपदः	U	उपमानपूर्वपद-व्यधिकरण-बहुव्रीहिः	BvU
		बहुपद-बहुव्रीहिः	Bb
द्वन्द्वः		**केवल**	
समासभेदाः	**अनुबन्धनाम**	**समासभेदाः**	**अनुबन्धनाम**
इतरेतरयोग-द्वन्द्वः	Di	केवलसमासः	S
समाहार-द्वन्द्वः	Ds	**द्विरुक्तिः**	
एकशेषः	E	द्विरुक्तिः	d

परिशिष्टः - ब : समासविग्रह-विधायक-नियमाः

अव्ययीभावः	
1	<x-y>A1 => y{6} f{x} where f maps x to the noun with same semantic content. A function f needs to be defined.
2	<x-y>A2 => x{3} विपरीतं वृत्तम्
3	<x-y>A4 => x{6} y{6} समाहारः
4	<x-y>A5 => x'{1} y{1} यस्मिन् देशे x' y' इत्यनयोः समानलिङ्गं स्यात् इति।
5	<x-y>A6 => x{6} y{6} समाहारः यदि x=द्वि स्यात् तर्हि x एवं y इत्यनयोः प्रयोगः द्विवचने स्यात्।
6	<x-y>A7 => x{6} y
तत्पुरुषः	
7	<x-y>T_n => x{n} y $2 \leq n \leq 7$
8	<x-y>Tn => x{n} y

9	<x-y>Tds => x{6;ba} y{6;ba} समाहारः
10	<x-y-z>Tb => x{1} y{1} z{1}
	कर्मधारयः
11	<x-y>K1 => x{1} तत् y{1} च
12	<x-y>K2 => x{1} च y{1} च
13	<x-y>K3 => x{1} च असौ y{1} च
14	<x-y>K4 => x{1} इव y{1}
15	<x-y>K5 => x{1} y{1} इव
16	<x-y>K6 => x{1} एव y{1}
17	<x-y>K7 => x{1} इति y{1}
	बहुव्रीहिः
18	<x-y>Bsd => x{6} च y{6} च यदन्तरालम्
19	<x-y>Bsp => x{3} च y{3} च प्रहृत्य इदं युद्धं प्रवृत्तम्
20	<x-y>Bsg => x{7} y{7} गृहीत्वा इदं युद्धं प्रवृत्तम्
21	<x-y>Bsmn => x' y{1} यस्य
22	<x-y>Bss => x{1} वा y{1} यस्य
23	<x-y>Bsu => x{1} इव y{1} यस्याः
24	<x-y>Bv => x y{1} यस्य
25	<x-y>Bvs => y{6} x' ये सन्ति ते
26	<x-y>BvS => y{3} सह
27	<x-y>BvU => x{6} इव y यस्य
	द्वन्द्वः
28	<x-y>Di => x{1} च y{1} च
29	<x-y>Ds => x{1} च y{1} च
30	<x-y>S => y{1} x{1}
31	<x-y>d => x y