विषय	हिंदी
प्रश्नपत्र सं. एवं शीर्षक	**P5: भाषाविज्ञान**
इकाई सं. एवं शीर्षक	**M19: प्रमुख आधुनिक परवर्ती व्याकरण**
इकाई टैग	**HND_P5_M19**

निर्माता समूह	
प्रमुख अन्वेषक	**प्रो. गिरीश्वर मिश्र** कुलपति, महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय, वर्धा (महाराष्ट्र) 442001 ईमेल : misragirishwar@gmail.com
प्रश्नपत्र समन्वयक	**डॉ. उमाशंकर उपाध्याय** पूर्व प्रोफेसर, भाषा विद्यापीठ महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय, वर्धा (महाराष्ट्र) 442001 ईमेल : usupadhyay@gmail.com
इकाई लेखक	**डॉ. धनजी प्रसाद** असिस्टेंट प्रोफेसर, भाषाविज्ञान एवं भाषा प्रौद्योगिकी विभाग, भाषा विद्यापीठ महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय, वर्धा (महाराष्ट्र) 442001 ईमेल : dhpr.langtech@gmail.com
इकाई समीक्षक	**डॉ. उमाशंकर उपाध्याय** पूर्व प्रोफेसर, भाषा विद्यापीठ महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय, वर्धा (महाराष्ट्र) 442001 ईमेल : usupadhyay@gmail.com
भाषा संपादक	**प्रो. गिरीश्वर मिश्र** कुलपति, महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय, वर्धा (महाराष्ट्र) 442001 ईमेल : misragirishwar@gmail.com

पाठ का प्रारूप

1. पाठ का उद्देश्य
2. प्रस्तावना
3. प्रमुख आधुनिक परवर्ती व्याकरण
4. निष्कर्ष

1. पाठ का उद्देश्य

इस पाठ के अध्ययन के उपरांत आप -

- शब्दकोशीय/शब्दवृत्तिक प्रकार्यात्मक व्याकरण (LFG) से परिचित हो सकेंगे।
- सामान्यीकृत पदबंध संरचना व्याकरण (GPSG) को समझ सकेंगे।

- शीर्ष-आधारित पदबंध संरचना व्याकरण (HPSG) को समझ सकेंगे।
- निर्भरता व्याकरण (DG) का परिचय पा सकेंगे।
- वृक्ष संलग्नक व्याकरण (TAG) से परिचित हो सकेंगे।

2. प्रस्तावना

चॉम्स्की द्वारा प्रस्तावित रचनांतरणपरक प्रजनक व्याकरण (1957) को भाषाविज्ञान के क्षेत्र में क्रांतिकारी व्याकरण माना गया है। इस व्याकरण में भाषा संबंधी नियमों को गणितीय नियमों की तरह प्रस्तुत करने पर बल दिया गया है।

इसके प्रतिपादन से भाषिक विश्लेषण और व्याकरण निर्माण का प्रारूप बदलने लगा। अब भाषावैज्ञानिक विश्लेषण वस्तुनिष्ठ (Objective) रूप से करके इसे भाषिक इकाइयों (Linguistic units) और तार्किक रचनाओं (Logical structures) में ढालकर प्रस्तुत किया जाने लगा। इस प्रकार प्रस्तावित व्याकरणों का उद्देश्य किसी भाषा विशेष का व्याकरण न होकर किसी भी भाषा के विश्लेषण हेतु नियमों के साँचे (Frames) प्रदान करना है। इसी कारण इन्हें व्याकरणिक फ्रेमवर्क (Grammatical framework) भी कहा गया है। 1970 के बाद से इस प्रकार के व्याकरण तेजी से उभरकर सामने आए हैं। इनकी संख्या 50 से भी अधिक है, किंतु सभी अधिक प्रचलित नहीं है। इनमें से प्रमुख पाँच को आगे दिया जा रहा है।

3. प्रमुख आधुनिक परवर्ती व्याकरण

'परवर्ती' से यहाँ तात्पर्य चॉम्स्की के रचनांतरणपरक प्रजनक व्याकरण को छोड़कर अन्य व्याकरणों से है। इस प्रकार के व्याकरण चॉम्स्की के व्याकरण को आधार बनाकर या उससे अलग किसी अन्य (जैसे- प्रकार्यात्मक) विधि को आधार बनाकर लिखे गए हैं। इनमें प्रमुख पाँच इस प्रकार हैं-

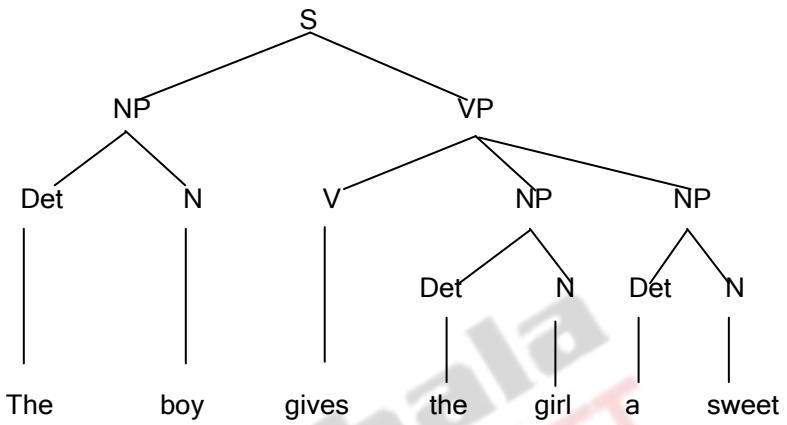
3.1 शब्दकोशीय प्रकार्यात्मक व्याकरण (Lexical Functional Grammar : LFG)

इसे 'कोशीय प्रकार्यात्मक व्याकरण' या 'शब्दवृत्तिक प्रकार्यात्मक व्याकरण' भी कहा गया है। इस सिद्धांत का विकास जोऑन ब्रेसनैन (Joan Bresnan) और रोनाल्ड कैपलैन (R. Kaplan) द्वारा 1970 के दशक के अंत में किया गया है। इसका विकास करते हुए ब्रेसनैन और कैपलैन ने प्राकृतिक भाषाओं के वाक्यों के विश्लेषण हेतु संगणकीय फार्मालिज्म (computational formalism) प्रदान करने का प्रयास किया है। 'Lexical Functional Grammar' (LFG) नाम का प्रयोग सर्वप्रथम ब्रेसनैन द्वारा संपादित 'The Mental Representation of Grammatical Relations' (1982) में किया गया।

इस व्याकरण में दो संरचनाओं को केंद्र में रखा गया है : सी-संरचना (C-structure) और एफ-संरचना (F-structure)। इन्हें संक्षेप में इस प्रकार देख सकते हैं-

(क) **सी-संरचना (C-structure):** सी-संरचना शब्द 'घटक संरचना' (Constituent structure) का संक्षिप्त रूप है। 'घटक' वे भाषिक इकाइयाँ हैं जिनसे मिलकर कोई वाक्य बना होता है, जैसे- शब्द, वाक्य और पदबंध। सी-

संरचना निरूपण में इन्हें ही इनकी वाक्यात्मक और शाब्दिक कोटियों (Syntactic and Word Categories) के रैखिक क्रम और अधिक्रमिक स्तर पर प्रदर्शित किया जाता है। इसके लिए एक संदर्भ-मुक्त व्याकरण होता है। यह नियमों का पदबंध संरचना वृक्षों (Phrase Structure Trees) के रूप में प्रतिरूपण है। इसमें शब्दों के क्रम और अधिक्रम दोनों को प्रस्तुत किया जाता है। इसे 'The boy gives the girl a sweet' वाक्य के विश्लेषण में इस प्रकार देखा जा सकता है-



आरेख : 01

(ख) एफ-संरचना (F-Structure): इसका पूरा नाम प्रकार्यात्मक संरचना (Functional structure) है। इसके माध्यम से बाह्य स्तर (Surface level) पर व्याकरणिक प्रकार्यों को निरूपित किया जाता है। इसमें संरचनात्मक और कोशीय सूचनाएँ संग्रहीत होती हैं। LFG में प्रत्येक शब्द को कुछ लक्षण (Feature) प्रदान किए जाते हैं, जिन्हें ही गुण (Attribute) कहा जाता है। इनके संभव मूल्य (Value) व्याकरण द्वारा निर्धारित होते हैं। वाक्य निर्माण के लिए शब्दों के साथ जुड़े गुणों (Attributes) और उनके मूल्यों (Values) का सहसंबंधित होना आवश्यक है। इस कारण एफ-संरचना को गुण मूल्यों (Attribute values) का संरचित रूप भी कहा गया है। इसके अंतर्गत सूचनाओं के निरूपण को तकनीकी रूप से 'गुण-मूल्य आधात्रियाँ' (Attribute value matrices) कहा जाता है। कोई भी एफ-संरचना गुण-मूल्य युग्मों (Attribute value pairs) का सीमित समुच्चय होती है जिसमें गुण एक प्रतीक होता है और इसके कुछ मूल्य निर्धारित होते हैं। इसे उपर्युक्त वाक्य के ही आधार पर शब्दों के साथ इस प्रकार देखा जा सकता है-

$$\begin{bmatrix} \text{PRED} & \text{'give}\uparrow\text{<SUBJ}\uparrow,\ \text{OBJ2}\uparrow,\ \text{OBJ}\uparrow\text{>'} \\ \text{SUBJ} & \begin{bmatrix} \text{SPEC} & \text{the} \\ \text{PRED} & \text{'boy'} \end{bmatrix} \\ \text{OBJ2} & \begin{bmatrix} \text{SPEC} & \text{the} \\ \text{PRED} & \text{'girl'} \end{bmatrix} \\ \textbf{OBJ} & \begin{bmatrix} \textbf{SPEC} & \text{a} \\ \textbf{PRED} & \text{'sweet'} \end{bmatrix} \\ \textbf{TENSE} & \text{PRESENT} \\ \textbf{ASPECT} & \text{INDEFINITE} \end{bmatrix}$$

आरेख : 02

एफ़-संरचनाओं पर सुनिर्मितता शर्तें (Well-formedness Conditions on F-Structures)

वाक्य निर्माण के लिए शब्दकोशीय प्रकार्यात्मक व्याकरण में कुछ सुनिर्मितता शर्तें दी गई हैं। इनमें से तीन प्रमुख निम्नलिखित हैं -

(1) **अद्वितीयता (Uniqueness) :** इसके अनुसार प्रत्येक एफ-संरचना में एक गुण के लिए एक ही मूल्य होना चाहिए। दो वाक्यात्मक घटकों के एकीकरण (Unification) के लिए उनमें एक गुण के लिए समान मूल्य का होना आवश्यक है। जैसे -

$$\begin{pmatrix} \text{NUM} & \text{Plural} \\ \text{PRED} & \text{men} \end{pmatrix} \qquad \begin{pmatrix} \text{NUM} & \text{Singular} \\ \text{SPEC} & \text{a} \end{pmatrix}$$

बहुवचन संज्ञा के साथ एकवचन विशेषक नहीं जोड़ा जा सकता।

(2) **संगति (Coherence) :** यह अर्थ की दृष्टि से दो शब्दों (घटकों) के परस्पर जुड़ने से संबंधित है। जैसे - वह आग से पानी सींच रहा है।

वाक्य नहीं बन सकता। क्योंकि सींचने का करण 'आग' और लक्ष्य 'पानी' अर्थ की दृष्टि से संबद्ध नहीं होते।

(3) **पूर्णता (Completeness)** : किसी विधेय रचना में जितने भी गुण (Attribute) दिए गए हों उनके लिए वाक्य में मूल्य होना चाहिए, जैसे - उपर्युक्त अंग्रेजी वाक्य की विधेय रचना इस प्रकार है -

'give <SUBJ, OBJ2, OBJ>'

इसके अनुरूप वाक्य में दो कर्म या प्रत्यक्ष कर्म न हो तो वाक्य अपूर्ण होगा:

*the boy gives.

*the boy gives the girl.

3.2 सामान्यीकृत पदबंध संरचना व्याकरण (Generalised Phrase Structure Grammar : GPSG)

इस व्याकरण का विकास 1970 तथा 1980 के दशक में जेराल्ड गजदर (Gerald Gazdar) और उनके सहयोगियों द्वारा किया गया, जिनमें क्लैन (Ewan Klain), सैग (Ivan Sag) और पलम (Geoffrey Pullum) आदि प्रमुख हैं। सन् 1985 में इन विद्वानों की 'Generalized Phrase Structure Grammar' नामक पुस्तक प्रकाशित हुई जो इस व्याकरण का आधार ग्रंथ है।

GPSG एक संदर्भमुक्त व्याकरण है, जिसके द्वारा भाषाओं की वाक्यात्मक और आर्थी व्यवस्था का विश्लेषण किया जाता हैइस व्याकरण में कोई भी कोटि लक्षण विशिष्टीकरणों (Feature specifications) का समुच्चय होती है जिसे नियमों द्वारा प्राप्त (Access) किया जा सकता है।

GPSG के अंतर्गत प्रयुक्त होने वाले कुछ प्रमुख नियम और घटक निम्नलिखित हैं:

(1) **सन्निकट प्रभुत्व (ID- Immediate Dominance) नियम** :ये नियम विभिन्न कोटियों के बीच सन्निकट प्रभुत्व के प्रतिबंधों को प्रदर्शित करते हैं। सन्निकट प्रभुत्व एकाधिक घटकों के बीच अधिक्रमिक संबंधो को स्थापित करने का कार्य करते हैं जिससे कि पदबंध संरचनाओं का निरूपण किया जा सके।

(2) **अधि-नियम (Metarules):** ये नियम सन्निकट प्रभुत्व (ID) नियमों की कोटियों के संबंध में सामान्यीकरणों (Generalizations) को प्रदर्शित करते हैं। इन नियमों का इनपुट ID नियम pattern निश्चित होता है जिसमें एक मातृ (Mother) कोटि, एक अंगजा (Daughter) कोटि और एक बहुसमुच्चय चर होते हैं।

(3) **रैखिक पूर्वगमन (LP- Linear Precedence) नियम** : रैखिक पूर्वगमन का अर्थ है, एक रैखिक स्तर पर घटकों का क्रम में एक के बाद एक आना। वाक्य से शब्द तक के स्तर पर कौन-सा घटक पहले आएगा और कौन-सा बाद में, को ये निर्धारित करते हैं। इन नियमों द्वारा एकाधिक घटकों के बीच संबंधो को स्थापित किया जाता है जिससे पदबंध संरचनाओं का निर्माण किया जा सके।

(4) **लक्षण सहघटना प्रतिबंध (FCR- Feature Cooccurance Restrictions):** कौन-से शब्द के साथ कौन-सा शब्द आएगा, यह दोनों की कोटियों और लक्षणों के आधार पर तय होता है। इनके द्वारा केवल अनुमेय कोटियों (Permissible categories) के साथ जुड़ने वाली लक्षण संरचनाओं को सहसंबंधित किया जाता है।

(5) **लक्षण विशिष्टीकरण वितथ (FSD- Feature Specification Defaults)** : इनके द्वारा अविशेषीकृत लक्षणों के लिए मूल्य प्रदान किए जाते हैं।

(6) **सार्वभौमिक लक्षण दृष्टांतीकरण (UFI- Universal Feature Instantiation):** ये वृक्षों में नोडों के अनुमेय समीपीय समुच्चयों को प्रतिबंधित करते हैं। इन लक्षण नियमों के अंतर्गत शीर्ष लक्षण रूढ़ि (HFC- Head Feature Convention), चरण लक्षण नियम (FFP- Foot Feature Principle) और नियंत्रण अन्विति नियम (CAP- Control Agreement Principle) आदि आते हैं।

GPSG में इन नियमों और प्रतिबंधों का प्रयोग क्रमानुसार किया जाता है।

GPSG में नियम

GPSG में प्रयुक्त होने वाले नियमों के साथ दो प्रकार की सूचनाएँ होती हैं: वाक्यात्मक और आर्थी। उदाहरणस्वरूप निम्नलिखित संदर्भ-मुक्त नियम को देखें:

S → NP, VP

इसका तात्पर्य है कि कहीं भी प्रतीक 'S' तो उसे 'NP, VP' के रूप में पुनर्लेखित किया जा सकता है। अर्थात 'NP, VP' के समुच्चय द्वारा इसे प्रतिस्थापित किया जाए जिसमें 'NP' 'VP' से पहले है। इन दोनों बातों को अलग-अलग नियमों में भी रखा जा सकता है जिसमें इनका स्वरूप इस प्रकार होगा:

S → NP, VP (1)

NP << VP (2)

1. सन्निकट प्रभुत्व (ID) नियम = प्रतीकों के समुच्चय {S} को प्रतीकों के समुच्चय 'NP, VP' द्वारा प्रतिस्थापित किया जा सकता है।
2. रैखिक पूर्वगमन (LP) नियम = कोई भी सन्निकट प्रभुत्व नियम जिसमें 'NP' और 'VP' आए हों, 'NP' 'VP' से पहले होगा।

बाद में विकसित इसके संशोधित रूप को 'संशोधित सामान्यीकृत पदबंध संरचना व्याकरण' (Revised Generalized Phrase Structure Grammar : RGPSG) नाम दिया गया है।

3.3 शीर्ष-चालित पदबंध संरचना व्याकरण (Head-Driven Phrase Structure Grammar : HPSG)

इसका HDPSG और HPSG दोनों रूपों में संक्षिप्तीकरण किया जाता है, किंतु HPSG ही अधिक प्रचलित है। इसका विकास कार्ल पोलार्ड (Carl Polard) और इवैन सैग (Ivan Sag) द्वारा 1980 के दशक में किया गया। यह एक प्रतिबंध-आधारित, शब्दवृत्तिक (Lexicalized) उपागम है। इसमें मानव भाषाओं की व्यवस्था की मॉडलिंग हेतु प्रकारित लक्षण संरचनाओं (Typed Feature Structure) की मुख्य भूमिका होती है। इस व्याकरण को अव्युत्पादक प्रजनक व्याकरण कहा गया है। HPSG पर सबसे अधिक प्रभाव GPSG का है। कुछ विद्वानों द्वारा इसे GPSG का उत्तराधिकारी (Successor) कहा गया है। इस व्याकरण के विकासकर्ताओं ने भाषाविज्ञान और व्याकरण के अतिरिक्त अन्य क्षेत्रों से भी कुछ संकल्पनाओं को लिया है; यथा: संगणकविज्ञान से डाटा टाइप सिद्धांत, ज्ञान प्रतिरूपण एवं सस्यूर की 'प्रतीक' (sign) की संकल्पना आदि।

HPSG की आधारभूत संकल्पनाएँ (Basic Concepts of HPSG):

(1) **लक्षण संरचना (Feature Structure)** : जैसा कि उपर भी दोनों व्याकरणों में संकेत किया जा चुका है, लक्षण संरचना गुण-मुल्य युग्मों (Attribute-Value pairs) का समुच्चय होती है। इसे 'directed acyclic graph' (DAG) के रूप में भी प्रदर्शित किया जा सकता है, जिसमें चरों के मूल्य को सहसंबंधित करने वाले नोड एवं चरों के नाम हेतु मार्ग (Paths) होंगे। लक्षण संरचनाओं पर कुछ संक्रियाएँ भी की जाती हैं; जैसे: एकीकरण (Unification) आदि।

(2) **गुण मूल्य आधात्री (Attribute Value Matrix: AVM)** : HPSG में गुण-मूल्य आधात्री वह इकाई है जिसका प्रयोग लक्षण संरचनाओं के वर्ग की व्याकरणिक विशेषताओं को प्रदर्शित करने के लिए किया जाता है, जिसमें लक्षण नाम को बाईं तरफ एक कॉलम में और मूल्य दाईं तरफ लिखा जाता है। आधात्री में दो कॉलम होते हैं : एक 'लक्षण नाम' के लिए और दूसरा 'मूल्य'के लिए। उदादरण के लिए एक AVM के स्वरूप को नीचे दिखाया जा रहा है:

$$\begin{bmatrix} \textit{category} & \textit{noun phrase} \\ \text{agreement} & \begin{bmatrix} \textbf{number} & \textit{singular} \\ \textbf{person} & \textit{third} \end{bmatrix} \end{bmatrix}$$

इस गुण-मूल्य आधात्री में दो लक्षण हैं :'Category' और 'Agreement'. इनमें 'category' का मूल्य 'noun phrase' है; किंतु 'agreement' के मूल्य को एक दूसरी लक्षण संरचना में निरूपित किया गया है जिसके अंदर दो लक्षण 'number' और 'person' हैं। इनमें 'number' का मूल्य 'singular' और 'person' का मूल्य 'third' है।

(3) **एकीकरण (Unification) :** HPSG में एक से अधिक लक्षण संरचनाओं को जोड़ने अथवा परस्पर सह-संबंधित करने की प्रक्रिया को एकीकरण कहते हैं। दो शब्दों में एक प्रकार के मूल्यों के होने पर ही उनका एकीकरण किया जाता है, जैसे- एकवचन संज्ञा के साथ एकवचन क्रियारूप ही आ सकता है।

(4) **शीर्ष (Head) :** 'शीर्ष' इस व्याकरण की केंद्रीय इकाई है। किसी पदबंध के जिस घटक पर दूसरे घटक आश्रित होते हैं, वह उस पदबंध का शीर्ष होता है। इसकी मुख्य भूमिका पदबंध एवं वाक्य स्तर पर होती है; जहाँ लक्षणों एवं मूल्यों का परीक्षण आदि करते हुए एकीकरण की क्रिया होती है। किसी भी पैरेंट नोड के लिए अधिकांश

लक्षण उसके चिल्ड्रेन नोडों में से प्रमुख नोड अर्थात 'शीर्ष' से लिया जाता है। जैसे- किसी संज्ञा पदबंध के लक्षण निर्धारण उसके 'संज्ञा पद' से होगा।

इस प्रकार विभिन्न घटकों, नियमों और प्रतिबंधों की व्यवस्था द्वारा HPSG के व्याकरणिक ढाँचे का गठन होता है और इसके द्वारा भाषिक विश्लेषण की पृष्ठभूमि तैयार होती है। इस व्याकरण में एक केंद्रीय भाग इसका शब्दकोशीय गठन (Lexical organization) है, जिसमें प्रत्येक शब्द के साथ उसकी गुण-मूल्य आधात्री में मूल्यों को दिया जाता है। दूसरा भाग पदबंधीय एवं वाक्यात्मक संरचना (Phrasal and Syntactic Structure) का है। इसमें सामान्यतः पदबंध संरचना नियमों के बजाय भाषिक तत्वों के बीच प्रभुत्व संबंधों (Dominance relations) को लक्षण संरचनाओं के साथ स्थान दिया जाता है।

3.4. निर्भरता व्याकरण (Dependency Grammar : DG)

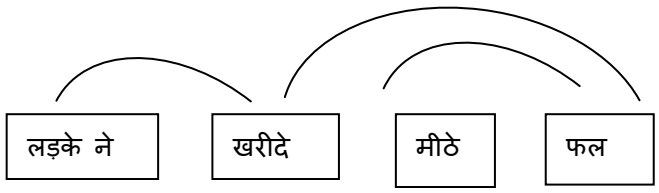
वैसे तो आधुनिक भाषाविज्ञान में निर्भरता व्याकरण को 1950 के दशक से स्थान मिलता है किंतु इसकी ऐतिहासिक पृष्ठभूमि बहुत प्राचीन है जो भारतीय परिप्रेक्ष्य में संस्कृतकाल तक जाती है। इस व्याकरण के आधुनिक रूपात्मक (formal) स्वरूप का आरंभ 1950 के दशक में Lucian Tensnièr के कार्यों से हुआ। Tensnièr एक फ्रांसीसी प्रोफेसर थे जो एक बहुभाषी भी थे। निर्भरता व्याकरण से संबंधित उनकी कृति 'Elèments de Syntaxe structurale' 1959 में प्रकाशित हुई जबकि उनकी मृत्यु 1954 में ही हो चुकी थी। 1960 के दशक में अनेक विद्वानों द्वारा निर्भरता आधारित कार्य आरंभ किए गए जो मुख्यत: वाक्यविन्यास (Syntax) से संबंधित हैं।

निर्भरता संरचना (Dependency Structure)

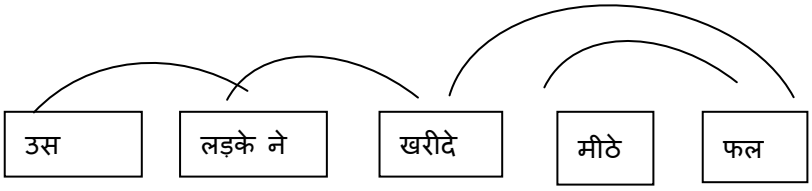
इसमें किसी भी वाक्य की वाक्यात्मक संरचना उसमें आए हुए घटक शब्दों के बीच द्विआधारी (Binary) संरचनात्मक संबंधों पर आधारित होती है। यह संबंध दो घटकों के बीच परस्पर निर्भरता का होता है। अर्थात् दो घटकों में एक मूल या शीर्ष होता है और दूसरा उस पर निर्भर होता है। उदाहरण के लिए निम्नलिखित आरेख में विभिन्न घटकों के बीच संबधों को देखें :

इस आरेख में दिखाया गया है कि 'A' और 'D' दोनों घटक 'B' पर निर्भर हैं, जबकि 'C' 'D' पर निर्भर है। इसे एक उदाहरण द्वारा समझने के लिए निम्नलिखित वाक्य को देखें:

लड़के ने मीठे फल खरीदे।

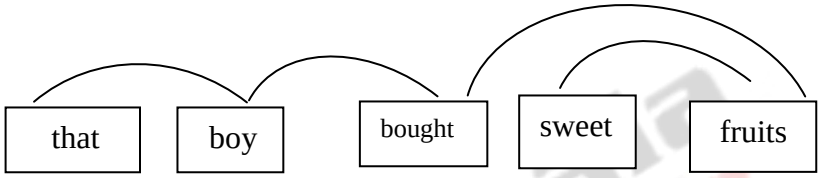


इस वाक्य में यदि लड़के शब्द से पूर्व एक निर्देशक 'उस' जोड़ दिया जाए तो यह घटक 'लड़का' पर निर्भर होगा और आरेख इस प्रकार होगा:



(नोट : इस उदाहरण में आरेख निर्माण में सुविधा के लिए क्रिया को बीच में रखा गया है। इसे अंत में भी रखा जा सकता है।)

इसी वाक्य का अंग्रेजी रूपांतरण करते हुए आरेख नीचे दिया जा रहा है:



निर्भरता व्याकरण (DG) को आधार बनाते हुए 1960 के दशक से अब तक भिन्न-भिन्न रूपों में अनेक व्याकरणिक ढाँचों का विकास किया गया है। उनमें से प्रमुख पाँच निम्नलिखित हैं:

1. प्रकार्यात्मक प्रजनक विवेचन (Functional Generative Description)
2. निर्भरता एकीकरण व्याकरण (Dependency Unification Grammar)
3. अर्थ पाठ सिद्धांत (Meaning Text Theory)
4. प्रकार्यात्मक निर्भरता व्याकरण (Functional Dependency Grammar)
5. शब्द व्याकरण (Word Grammar)

3.5 वृक्ष संलग्नक व्याकरण (Tree Adjoining Grammar : TAG)

वृक्ष संलग्नक व्याकरण को सामान्यत: इसके अंग्रेजी नाम के संक्षिप्त रूप 'TAG' (टैग) कहकर संबोधित किया जाता है। यह प्रो. अरविंद जोशी द्वारा विकसित किया गया एक व्याकरणिक फ्रेमवर्क है। यह एक प्रकार के संदर्भ -मुक्त व्याकरण की तरह ही है किंतु इसमें प्रतीकों की जगह वृक्षों का प्रयोग किया जाता है। इस व्याकरण में एक वृक्ष के नोडों से दूसरे उपयुक्त वृक्षों को मिलाते अथवा जोड़ते हुए नई-नई वृक्ष संरचनाओं का निर्माण किया जाता है जो भाषिक अभिव्यक्तियों का प्रतिनिधित्व करती हैं।

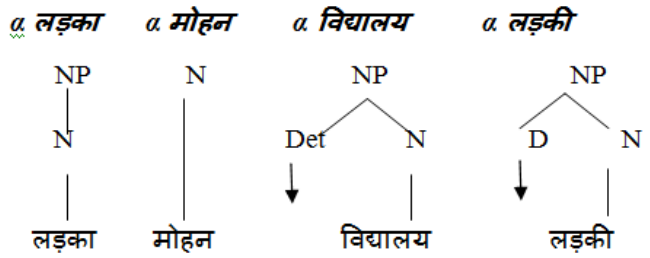
टैग में वृक्ष संरचना (Tree Structure in TAG)

टैग में आधारभूत रूप से दो प्रकार की वृक्ष संरचनाएँ होती हैं: आरंभिक वृक्ष (Initial Tree) और सहायक वृक्ष (Auxiliary Tree)। इन दोनों को मिलाकर समेकित रूप से मूल वृक्ष (Elementary Tree) कहा जाता है। अत: तकनीकी रूप से इस व्याकरण को 'I' (Initial) और 'A' (Auxiliary) वृक्षों का सीमित समुच्चय कहा जा सकता है। इन वृक्षों की संरचना को निम्नलिखित प्रकार से समझा जा सकता है:

(क) आरंभिक वृक्ष (Initial Tree)

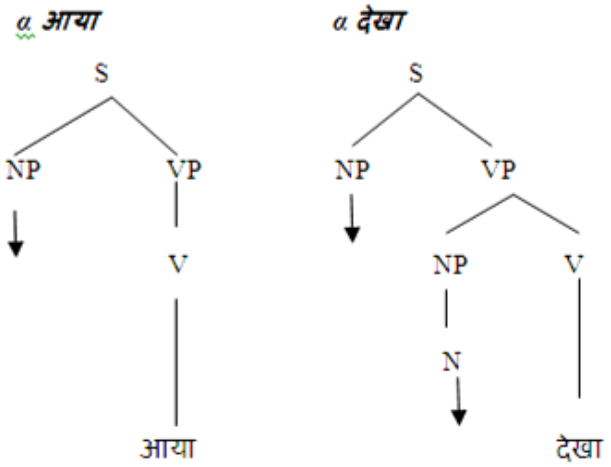
आरंभिक वृक्ष वे वृक्ष हैं जिनमें एक मूल नोड और एक समापक नोड अवश्य रहते हैं। इन वृक्षों के नोडों के दो प्रकार किए जाते हैं- समापक नोड (Terminal nodes) और असमापक नोड (Non-terminal nodes)। इन नोडों पर समापक और असमापक प्रतीकों का प्रयोग किया जाता है। इन वृक्षों को एक विशेष अक्षर 'α' द्वारा प्रदर्शित किया जाता है। कुछ α वृक्षों के उदाहरण निम्नलिखित हैं:

संज्ञाओं के α वृक्ष:

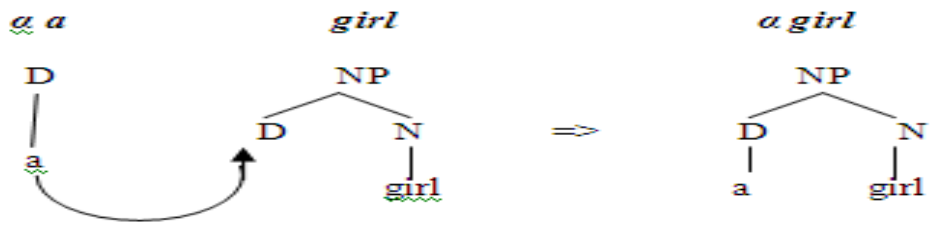


क्रियाओं के α वृक्ष

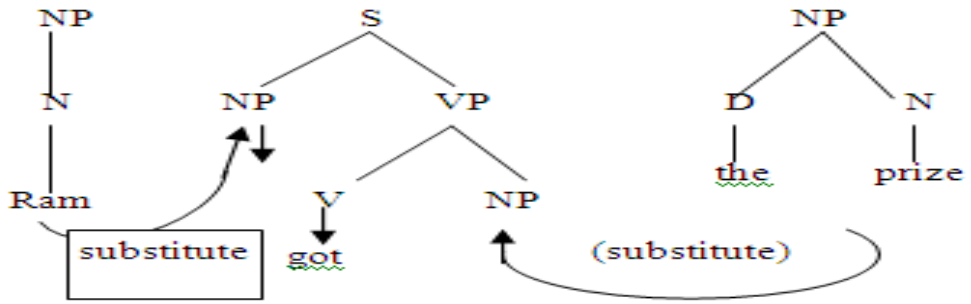
उपर्युक्त चित्रों में दिए गए सभी वृक्ष α वृक्षों के उदाहरण हैं। इनमें देखा जा सकता है कि क्रिया अपनी वाक्यात्मक संरचना के साथ आ रही है और अन्य घटक अपनी पदबंधीय संरचना के साथ। इन वृक्षों में दो प्रकार के नोड हैं- असमापक नोड और समापक नोड। इनमें सभी लीफ नोडों के साथ उनके प्रतीकों को रखा गया है। जिन लीफ नोडों के साथ निम्नगामी तीर (Down arrow) का प्रयोग किया गया है वे 'प्रतिस्थापन नोड' (Substitution nodes) कहलाते हैं। प्रतिस्थापन नोडों पर प्रतिस्थापन की प्रक्रिया होती है जिसके अंतर्गत आरंभिक वृक्षों में इन नोडों को दूसरे वृक्षों द्वारा प्रतिस्थापित करके नए वृक्षों का व्युत्पादन किया जाता है। उदाहरण के लिए अंग्रेजी के निम्नलिखित वृक्षों को देखें:



पदबंध स्तर



वाक्य स्तर:

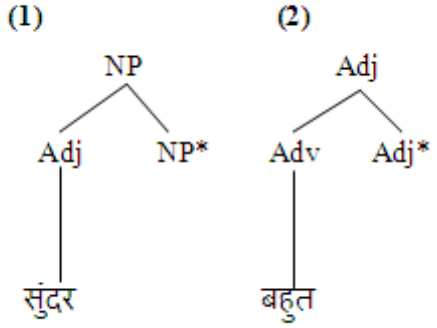


इस प्रकार से निर्मित होने वाले वृक्षों को व्युत्पादित वृक्ष कहते हैं। यहाँ पर एक महात्वपूर्ण बात यह है कि किसी भी वृक्ष में नोडों के पते (Addresses) दिए जा सकते हैं। अर्थात् उन्हें संख्यात्मक मूल्य (0, 1, 2 आदि) प्रदान करते हुए चिह्नित किया जा सकता है जिससे कि किसी भी आरंभिक अथवा व्युत्पादित वृक्ष में नोड, उस पर के प्रतीक अथवा उस पर होने वाली क्रिया को इंगित या संकेतित किया जा सके।

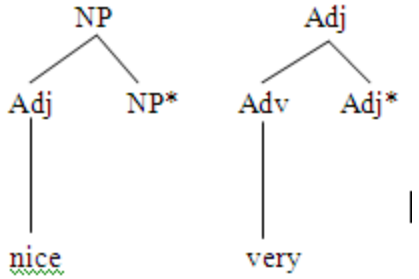
(ख) सहायक वृक्ष (Auxiliary Tree)

ये ऐसे वृक्ष होते हैं जिनका मूल नोड और एक फुट नोड के प्रतीक समान होते हैं। अत: इन दोनों नोडों पर असमापक प्रतीक होते हैं। समापक प्रतीकों का प्रयोग इनके अलावा दूसरे नोडों पर किया जाता है। इन वृक्षों का प्रयोग संलग्नन (adjoining) की प्रक्रिया में किया जाता है। इन वृक्षों को 'β' द्वारा प्रदर्शित करते हैं। कुछ β वृक्षों के उदाहरण निम्नलिखित हैं:

हिंदी



अंग्रेजी:



TAG में होने वाली प्रक्रियाएँ

वृक्ष संलग्नक व्याकरण में मुख्यत: दो प्रक्रियाएँ होती हैं- प्रतिस्थापन और संलग्नन। इन्हें विस्तृत रूप से इस प्रकार समझा जा सकता है:

1. प्रतिस्थापन (Substitution)

टैग के अंतर्गत की जाने वाली यह प्रथम मूल प्रक्रिया है। इसके माध्यम से दो आरंभिक (α) वृक्षों को जोड़कर नए वृक्ष का व्युत्पादन किया जाता है। आरंभिक वृक्षों में कुछ प्रतिस्थापन नोड होते हैं जिन पर दूसरे आरंभिक वृक्ष आकर जुड़ जाते हैं और एक नए वृक्ष का व्युत्पादन होता है।

2. संलग्नन (Adjoining)

वैसे तो प्रतिस्थापन द्वारा ही अनेक प्रकार के वाक्यों का व्युत्पादन किया जा सकता है किंतु अनेक स्तरों पर इसकी सीमाएँ प्राप्त होती हैं। अत: व्याकरण को अधिक शक्तिशाली एवं पूर्ण बनाने के लिए 'संलग्नन' (Adjoining) की संकल्पना दी गई है। इसे समझने के लिए सहायक वृक्षों (β) की संरचना पर ध्यान दें। उन वृक्षों के मूल नोड (r) और चरण नोड (f) पर एक ही समापक वृक्ष का प्रयोग किया जाता है। इनमें से चरण नोड (f) को '*' द्वारा चिह्नित किया जाता है। इसके अलावा अन्य सभी लीफ नोडों पर या तो समापक प्रतीक होते हैं या वे प्रतिस्थापन नोड होते हैं। किसी भी सहायक वृक्ष में कम से कम एक समापक प्रतीक (शब्दकोशीय इकाई) का होना आवश्यक है।

इस प्रकार यह व्याकरण वाक्य रचना हेतु आरंभिक वृक्षों और सहायक वृक्षों को जोड़ने के नियम देता है। इन वृक्षों को जोड़ने के लिए प्रतिस्थापन और संलग्नन की प्रक्रियाएँ होती हैं। इन प्रक्रियाओं के समय वृक्षों के साथ जुड़ी लक्षण संरचनाओं के एकीकरण का भी ध्यान रखा जाता है। इसी प्रकार जोशी द्वारा 'वृक्ष परिवारों' (tree families) की भी बात की गई है।

4. निष्कर्ष

संक्षेप में, आधुनिक परवर्ती व्याकरणों या व्याकरणिक फ्रेमवर्कों में नियमों को तार्किक रचनाओं के रूप में रखने पर बल दिया गया है। इन व्याकरणों में भाषिक इकाइयों को कोशीय संरचनाओं या घटकों में इस प्रकार रखने पर बल दिया गया है कि उनके साथ उनकी कुछ वाक्यात्मक (व्याकरणिक) और आर्थी विशेषताएँ भी आ जाएँ। इन्हें ही कहीं लक्षण संरचना तो कहीं गुण-मूल्य आधात्री आदि नाम दिए गए हैं। अत: कुछ नियम जहाँ छोटी भाषिक इकाइयों को मिलाकर बड़ी भाषिक इकाई (शब्द → पदबंध → वाक्य) के निर्माण से संबंधित हैं, वहीं कुछ नियम ऐसे भी हैं जो इन विशेषताओं के समान होने का ध्यान रखते हैं। इससे सही वाक्यों के निर्मित और विश्लेषित होने का प्रतिशत बढ़ा है। इस प्रकार के कई व्याकरण हैं, जिनमें से प्रमुख पाँच से ऊपर परिचित कराया गया।