

हिंदी के संगणकीय विश्लेषण के लिए औपचारिक व्याकरण का एनएलपी में प्रयोग

विवेक त्रिपाठी

मानवतावादी अध्ययन विभाग

भारतीय प्रौद्योगिकी संस्थान (बी.एच.यु), वाराणसी - 221005

sopan.tripathi@gmail.com

सार

प्रस्तुत शोध सामान्य बोलचाल वाली हिंदी पर औपचारिक व्याकरण (formal grammar) का प्रयोग करता है। इसका उद्देश्य हिंदी वाक्य विन्यास और हिंदी शब्दार्थ को समझने के लिए वैज्ञानिक आधार बनाना है। इसमें प्रस्तावित सभी ४ भाषारं तार्किक हैं अर्थात् सभी भाषाओं में अलग-अलग तर्क को सम्मति किया गया है। जैसे कि कथन तर्क (propositional logic), विधेय तर्क (predicate logic), प्रकार तर्क (type logic) और समयसापेक्ष तर्क (temporal logic).

मूल हिंदी के व्याकरण की प्रकृति का वर्णन करने के लिए कुछ वाक्य-विन्यास और शब्दार्थ-नियम प्रस्तावित किये गए हैं। हिंदी व्याकरण की जटिलताओं को समझने के लिए एक मूल नमूना L_{1H} नाम की भाषा बना कर समझाया गया है। L_{1H} एक कृत्रिम भाषा है जिसका ऐसा नामकरण इसकी आधार भाषाओं की वजह से किया गया है। L_{1H} में '1' कथन तर्क को और 'H' हिंदी भाषा का प्रतिनिधित्व करता है। इसके बाद L_{1H} का कई बार परिवर्धन किया गया है। परवर्धित भाषाएँ (L_{2H} , L_{3H} और L_{4H}) को विभिन्न वाक्य-श्रेणियों के हिसाब से रखा गया है।

वाक्य-श्रेणियों का कुछ उदाहरण ऐसा हो सकता है जैसे कि (१) राम चलता है, (२) बच्चे चलते हैं, (३) राम चलता है और बच्चे चले जाते हैं, (४) राम चला था और बच्चे चलेंगे आदि। इन सभी वाक्यों में एक समानता और एक अंतर है। समानता यह है कि ये सभी वाक्य 'घोषणात्मक प्रकृति' (affirmative nature) के हैं। अंतर यह है कि इन सभी वाक्यों की क्रिया-संरचना अलग-अलग है। इसलिए नंबर (१) जैसे वाक्य L_{1H} भाषा से, (२) जैसे वाक्य L_{2H} से, (३) जैसे वाक्य L_{3H} से और (४) जैसे वाक्य L_{4H} से समझे जा सकते हैं।

हमारा मुख्य सैद्धांतिक चिंतन 'तार्किक वाक्यविन्यास' और 'भाषाई वाक्यविन्यास' के बीच के संबंध को समझना है। सभी व्याकरणिक नियमों को हमने सॉफ्टवेयर के माध्यम से कार्यान्वित भी किया है। उम्मीद है कि इस अध्ययन से हिंदी के लिए एक नया भाषाई ढाँचा विकसित होगा।

१. प्रस्तावना

औपचारिक व्याकरण हिंदी को वाक्य-संरचना में मदद करता है। यह लेख एक आधुनिक दृष्टिकोण को पेश करता है जिससे हमें हिंदी की वाक्य संरचना एवं अर्थ संरचना को समझने का एक मौका प्राप्त होता है। भाषा विज्ञान की दृष्टि से वाक्य संरचना को 'सिंटेक्स (syntax)' और अर्थ संरचना को 'सिमेंटिक्स (semantics)' भी कहा जाता है। इस

पद्धति द्वारा विश्व की लगभग सभी भाषाओं में अनुसंधान एवं अध्ययन किया गया है। अतः अनेक विद्वानों में फर्डिनेंड डी सौस्युर, नोम चोम्स्की, रिचर्ड मॉटेम्यु, बारबरा पार्टी आदि का नाम विशेष रूप से आता है।

हिंदी में एक विस्तृत और यथासंभव 'सिंटेक्स और सिमेंटिक्स' को समझने के लिए एक पूर्ण ढांचा तैयार करना ही इस शोध का उद्देश्य है। परन्तु हम ऐसा करना ही क्यों चाहते हैं? इसको समझने के लिए मैं आप से कुछ उदाहरण साझा करना चाहता हूँ।

उदाहरण:

- राम चलता है।
- राम चला जाता है।
- राम चलता हुआ जाता है।
- राम चल करके खाने जाता है।
- राम चलते चलते जाता है।

इन सभी वाक्यों को हमारी आम व्याकरणिक दृष्टि (जिसमें संज्ञा, सर्वनाम आदि रूप में वाक्यांशों को समझा जाता है) से भी देखा जा सकता है। अन्य दृष्टि यह हो सकती है कि इन सभी वाक्यों को सिर्फ ($X + Y$) के रूप में तोड़ दिया जाए। अब यहाँ 'X' वाक्य का वह घटक हो सकता है जो हर वाक्य में एक जैसा है और 'Y' वह घटक जो हर वाक्य में बदल रहा है। अतः यहाँ $X = \text{'राम चल'}$ और $Y = \text{'ता है' / 'आ जाता है' / 'ता हुआ जाता है' / 'करके खाने जाता है' / 'ते चलते जाता है'}$ आदि हो सकता है। इसी तरह की वाक्य विन्यास की प्रक्रिया चारों परिवर्धित भाषाओं को बनाने में लगाई गई है। फिर कंप्यूटर को यह प्रक्रिया गणित द्वारा समझाई गयी है। इसे समझने के बाद कंप्यूटर के लिए ऐसे वाक्य या इस तरह के अन्य वाक्यों को बनाने की प्रणाली पता चली।

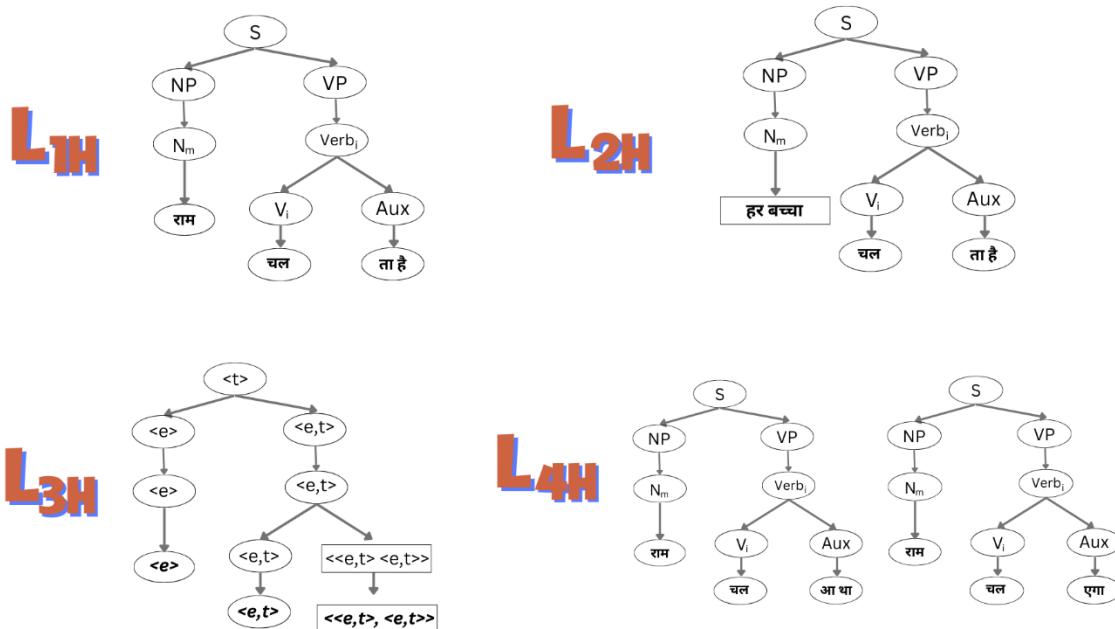
2. शोध-प्रविधि

X या Y घटक हिंदी वाक्यों में किस प्रकार होने चाहिए? - इसकी संकल्पना "सिंटेक्स के नियमों" द्वारा की जाती है। जैसे कि अगर मैं एक वाक्य जैसे कि "राम चलता है" का निर्माण करना चाहता हूँ, तो मैं यह कहूँगा कि "राम" एन.पी (NP) कैटेगरी का एक रूपिम है और "चलता है" वी.पी (VP) कैटेगरी का। यह एन.पी कैटेगरी हमारे लिए X के स्थान पे आ सकती है और वी.पी कैटेगरी Y के स्थान पर। यह संकल्पना आधुनिक काल में विश्व की कई भाषाओं के लिए की गयी है, जिसे हम "ट्रांसफॉर्मेशनल जेनेरेटिव ग्रामर [Transformational Generative Grammar (TGG)]" के नाम से जानते हैं। इस ग्रामर (अर्थात् व्याकरण) में सबसे बड़ा नाम "नोम चोम्स्की" का माना जाता है।

अब आते हैं अगली अवधारणा पर - क्योंकि हम हिंदी के वाक्य को किसी पूर्व-तय व्याकरणिक इकाई (जैसे कि संज्ञा, सर्वनाम आदि) के अनुसार नहीं तोड़ रहे हैं, बल्कि अपने स्व-सिद्धांत (कोई भी सिद्धांत जो आप अपने विमर्श में रख सकते हैं) के हिसाब से तोड़ रहे, इसलिए आप X और Y के गुण भी खुद से ही निर्धारित कर सकते हैं। यह भाषा को बनाने की या वाक्यों की संरचना को करने का बड़ा ही अनोखा व्यष्टिकोण है। इसी संकल्पना का प्रयोग करके इस शोध में हिंदी के वाक्यों की संरचना की गई है। उम्मीद है कि अब आपको "हिंदी अपनी वाक्य संरचना कैसे कर रही है?" इस प्रश्न का उत्तर भी धीरे-धीरे मिलने लगेगा।

जैसा कि मैंने कहा कि X और Y स्थापित व्याकरणिक इकाई नहीं हैं, इसलिए ये हिंदी व्याकरण के गुण के हिसाब से बंधे भी नहीं हैं। हम गुणों का निर्धारण स्वयं कर सकते हैं। हमने X और Y के लिए जो गुण निर्धारित किये हैं, वो "तर्क शास्त्र" विषय से आए हैं। आप चाहे तो किसी और विषय के गुण भी रख सकते हैं। गुण को निर्धारित करने के लिए हमारा प्रयोजन "तर्क-शास्त्र" इसलिए था क्योंकि हम चाहते थे कि वाक्य संरचना के बाद जो हिंदी वाक्य बने, उनसे हम तर्क भी कर सकें। जैसे कि अगर हमने कहा कि "राम चलता है" जगत में होने वाला एक सत्य वाक्य है, तो इसका अर्थ ये भी है कि "राम बैठा नहीं है" क्यूंकि उसके अन्दर चलने का गुण पहले से विद्यमान था। अतः "राम चलता है" का सत्य होना "राम बैठा नहीं है" इसको भी स्वयं सिद्ध कर देता है। इसी प्रकार यदि "राम चलता है" की सत्यता स्थापित है, तो यह ये भी निर्धारित कर सकता है कि "राम जाता है" क्योंकि जाने वाला व्यक्ति चलते हुए ही जाता है। इन्हीं गुणों को हमने अपने शोध में "सिमेंटिक्स" कहा है। इस प्रकार की संकल्पना का प्रयोग रिचर्ड मॉटेंग्यु द्वारा "मॉटेंग्यु ग्रामर [Montague Grammar (MG)]" के तहत किया गया है।

इन्हीं "सिमेंटिक्स और सिमेंटिक्स के नियमों" का हमने गूढ़ता से अध्ययन करके 4 प्रकार की भाषाएँ बनाई हैं, जिनके अन्दर वाक्य संरचना की शक्ति होने के साथ-साथ, अर्थ-निर्धारण के लिए तर्क का भी प्रयोग किया जा सकता है। ये 4 भाषाओं की संरचना कुछ इस प्रकार दिखाई देती हैं:



MG और TG क्रमशः 'द स्कूल ऑफ फिलोसोफी' और 'द स्कूल ऑफ लिंग्विस्टिक्स' के शोधकर्ताओं द्वारा अध्ययन के दो प्रमुख शास्त्रीय क्षेत्र रहे हैं। यह शोध हिंदी भाषा का विश्लेषण करने और उसके औपचारिक समकक्ष को विकसित करने पर केंद्रित है। हिंदी के लिए जो ढाँचा विकसित किया गया है वह MG और TG का मिश्रण है। सफल वाक्य-विन्यास और इस तरह की शब्दार्थ-प्रक्रिया से हिंदी के प्राकृतिक भाषा प्रसंस्करण [Natural Language Processing (NLP)] में आशाजनक योगदान देने के लिए यह एक प्रयास है।

जिन भाषा-विद्वानों को परिवर्धित भाषाओं के बीच का मूल अंतर जानना है, उनके लिए मैं यह कहना चाहता हूँ कि - ऊपर बनाई गई चारों परिवर्धित भाषाएँ अपनी तार्किक भाषाओं के समरूपी हैं। इस शोध में सभी परिवर्धित भाषाओं को एक सॉफ्टवेयर के माध्यम से प्रदर्शित भी किया गया है। जहाँ L_{1H} भाषा मुख्य रूप से संदर्भात्मक संज्ञा के लिए वाक्यात्मक नियम विकसित करने पर ध्यान केंद्रित करती है, वहाँ L_{2H} गैर-संदर्भात्मक संज्ञाओं के लिए परिमाणीकरण (जिसे अंग्रेजी में quantification कहा जाता है) का उपयोग करती है। L_{3H} भाषाई अभिव्यक्तियों को उनके शब्दार्थ-प्रकारों के आधार (जिन्हें Type कहा जाता है; मूल टाइप केवल दो प्रकार के होते हैं - 'इ टाइप' और 'टी टाइप') पर वर्गीकृत करती है, और L_{4H} हिंदी के भूत और भविष्य काल के लिए "समय अवधि" के तर्क का प्रयोग करती है।

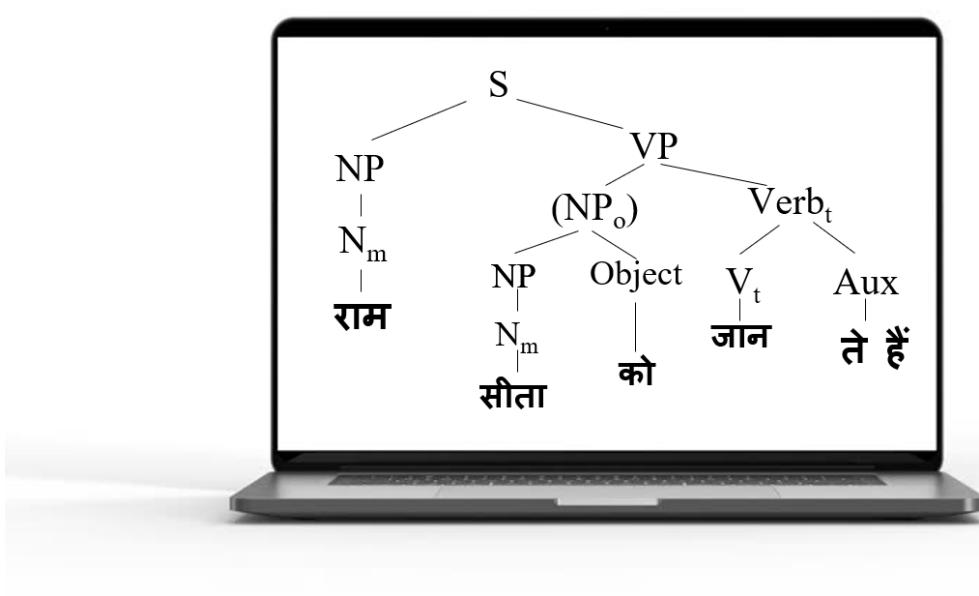
3. नवीनता एवं उपलब्धि

इस शोध अध्ययन से हिन्दी की वाक्य रचना एवं शब्दार्थ की कार्यप्रणाली की वैज्ञानिक समझ प्राप्त हुई है। परिवर्धित भाषाएँ जिनके भाषाई व्यवहार अलग-अलग हैं, वे मूल समाज में बोले जाने वाली हिंदी को वैज्ञानिक रूप से समझने के लिए एक संपूर्ण रूपरेखा प्रदान करते हैं। टाइप प्रणाली के प्रयोग से हिंदी के संयुक्त क्रियाओं के व्यवहार को समझने का एक वैकल्पिक दृष्टिकोण मिलता है। हिंदी में इस योगदान को कंप्यूटर हेतु क्रियान्वित करने के लिए हमने टी.एच.टी (THT) नामक पार्सर भी बनाया है जो हिंदी वाक्यों का भाषाई वाक्यविन्यास या उन्हें शाखाओं के रूप में तोड़ कर वाक्य को समझने का एक सरल तरीका बताता है। यह पार्सर छोटे या बड़े किसी भी हिंदी वाक्य को पार्स (वर्गीकृत) कर सकता है जो हिंदी के वाक्यविन्यास नियमों के अनुसार ही चलता है। हमने पाठकों की सुलभता हेतु सॉफ्टवेर से बनाये गए कुछ वाक्य-विन्यास वृक्ष या वाक्य-विन्यास वर्गीकरण (syntax tree) नीचे दिए हैं।

3.१ सिंटेक्स के नियमों का विवरण:

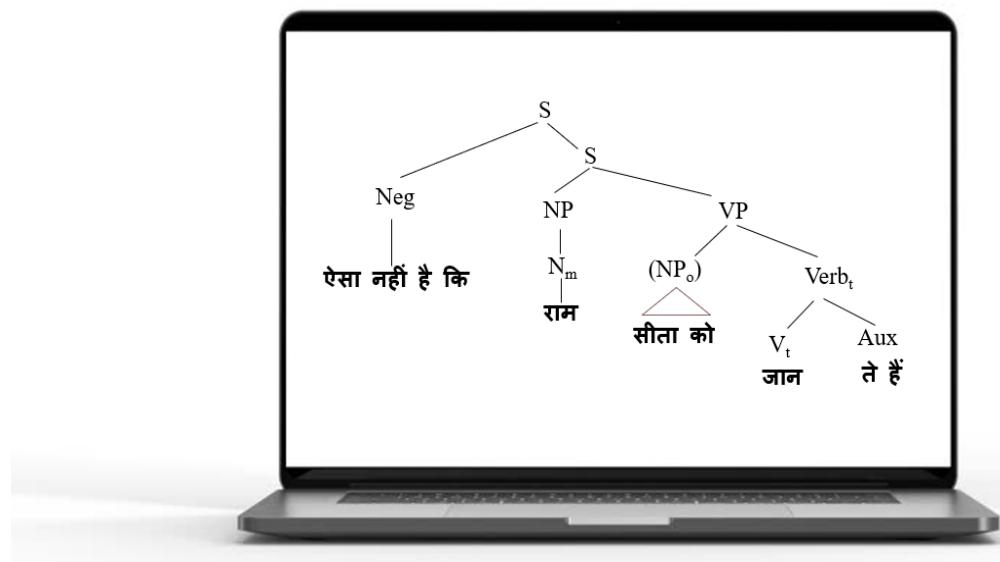
यह शोध इस लेख के माध्यम से 18 वाक्यात्मक नियमों और 15 अर्थ संबंधी नियमों को शामिल करता है। यह पेपर सहायक-क्रिया (ऑक्स) के लिए दो अलग-अलग उपचारों का प्रस्ताव करता है। ऑक्स वह भाषाई इकाई है जिससे क्रिया के होने का समय पता चलता है। ऑक्स का एक उदाहरण 'रहा है' है जहाँ 'रहा' क्रिया के किसी समय में लगातार चलने की स्थिति को बताता है और 'है' वर्तमान समय को दर्शाता है। पहले उपचार में 'ऑक्स' को एक वाक्य-विन्यास तत्व के रूप में माना जाता है और दूसरे उपचार में 'ऑक्स' को एक वाक्य-विन्यास श्रेणी के रूप में माना जाता है।

वाक्य १: राम सीता को जानते हैं |

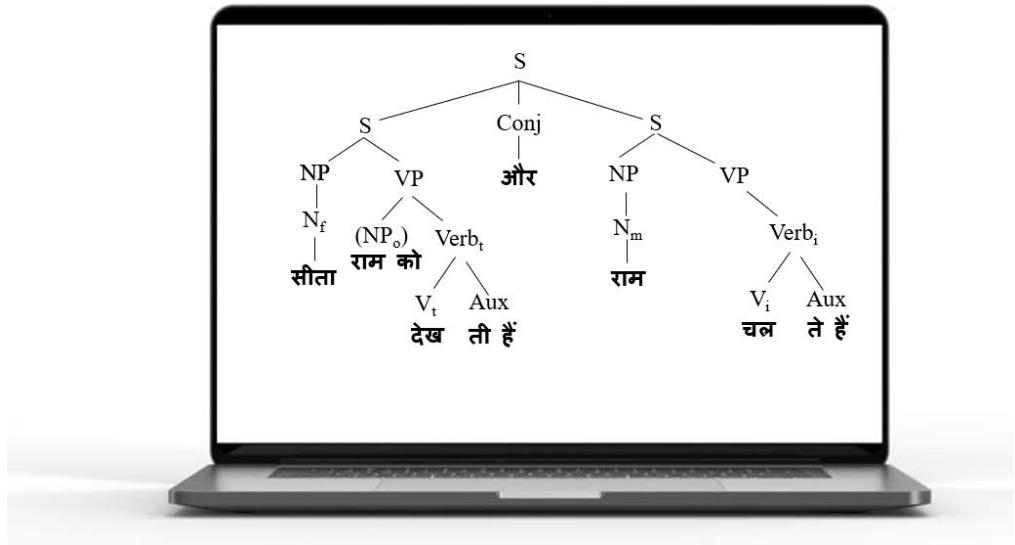


वाक्य २: राम सीता को नहीं जानते हैं |

(इस वाक्य को निम्नांकित रूप से बदल के तोड़ा गया है)

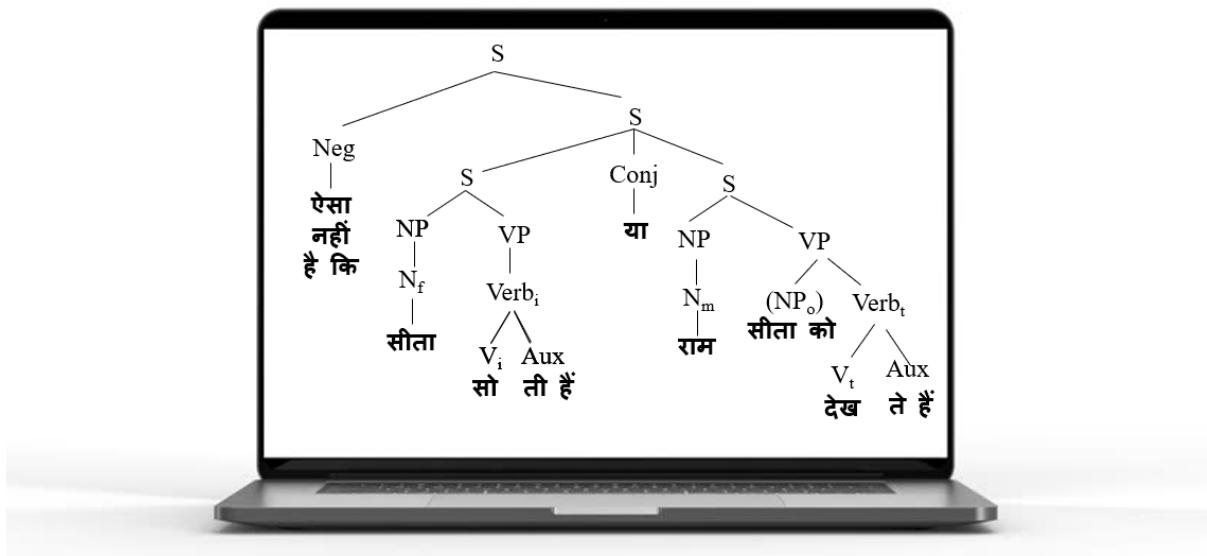


वाक्य ३: सीता राम को देखती हैं और राम चलते हैं |

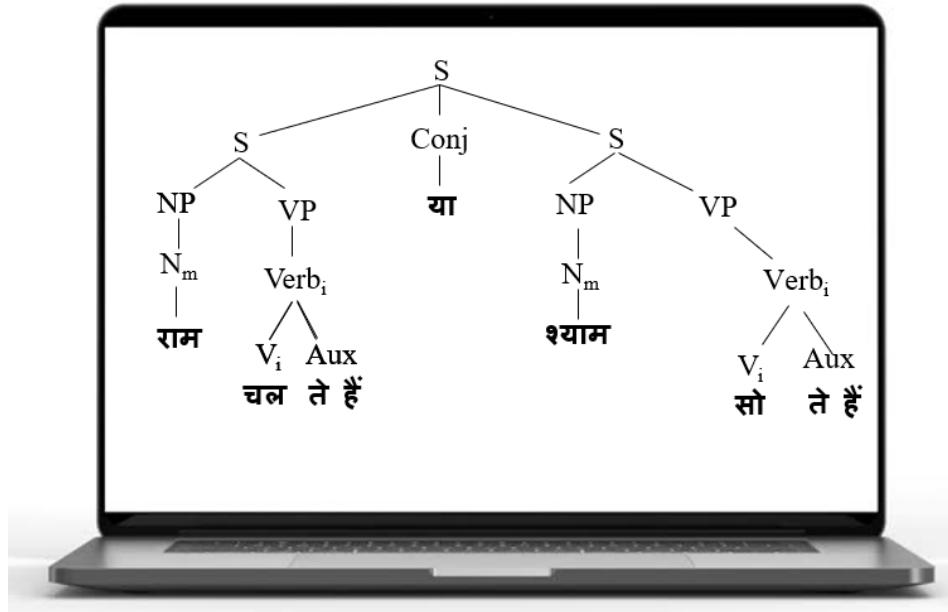


वाक्य ४: सीता नहीं सोती हैं या राम सीता को देखते हैं |

(ध्यान देने योग्य बातः इस वाक्य को निम्नलिखित रूप से बदला गया और फिर तोड़ा गया है) |



वाक्य ५: राम चलते हैं या श्याम सोते हैं ।

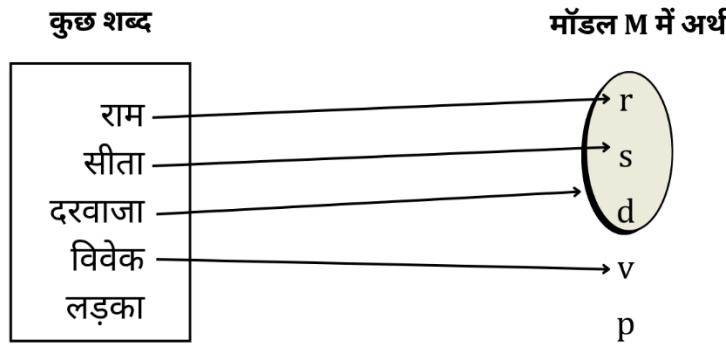


3.2 सिमेंटिक्स के नियमों का विवरण:

- अर्थ निकालने के लिए संकल्पना १

' α का अर्थ' दिखाने का तरीका है $[\alpha]$, जहां α एक भाषाई इकाई है । इसके अर्थ की गणना एक मॉडल M के अन्दर की जाती है । अतः अगर अलग-अलग मॉडल M बनाये जाए तो अनेकों अर्थ α को दिए जा सकते हैं । इसमें से कुछ अर्थ वास्तविक जगत में जैसे अर्थ माने जाते हैं वैसे भी हो सकते हैं और कुछ अर्थ भिन्न भी हो सकते हैं । उदाहरण के लिए अगर ' $\alpha = \text{गुलाब}$ ' है तो गुलाब को फूल न मानकर एक इंसान भी माना जा सकता है।

चलिए इस परिसंकल्पना से एक मॉडल M को बना कर देखते हैं जो मानव जगत के अर्थ और अन्य अलग-अलग अर्थों को संकलित करके बना है । M का एक उदाहरण निम्नलिखित है:



$\llbracket \text{राम दरवाजा है} \rrbracket = 1$

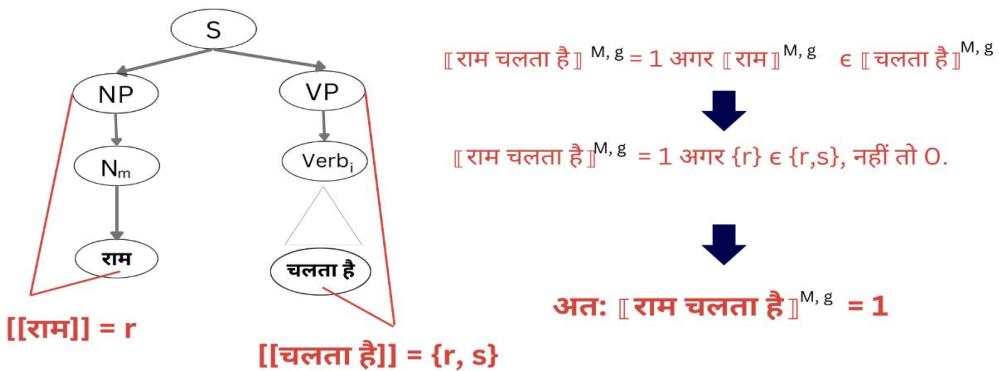
$\llbracket \text{राम लड़का है} \rrbracket = 0$

- अर्थ निकालने के लिए संकल्पना २

इस मॉडल M में, एक “व्याख्या फंक्शन g ” बनाया गया है जो राम, सीता, दरवाजा, लड़का आदि शब्दों को फंक्शन अनुसार अर्थ देता है, जैसे कि, राम का अर्थ है $\{r\}$, सीता का अर्थ है $\{s\}$, दरवाजा का अर्थ है $\{r, s, d\}$, और लड़का का अर्थ $\{\varphi\}$ है। इस मॉडल में लड़का का अर्थ $\{\varphi\}$ होने का मतलब है कि फंक्शन को नहीं पता कि इस शब्द से क्या अर्थ समझना है।

- अर्थ निकालने के लिए संकल्पना ३

प्रत्येक सुगठित संरचना (वह संरचना जो सिंटेक्स के नियम अनुसार बनी है, इसे हम wff कहते हैं) ही L_{1H} , L_{2H} , L_{3H} या L_{4H} के भाषा में शामिल हैं। हर वाक्य में भाषा की इकाई को या तो एक व्यक्ति के रूप में या एक विधेय के रूप में माना जाता है। विधेय एक फंक्शन के रूप में कार्य करते हैं और व्यक्ति इसी फंक्शन के आर्ग्युमेंट के रूप में। इस तरह के भ्रेद की पहचान सिंटेक्स के नियमों को देखने के बाद ही की जाती है। उदाहरण के लिए ‘राम चलता है’ में, $\llbracket \text{राम} \rrbracket$ का अर्थ $\{r\}$ और $\llbracket \text{चल} \rrbracket$ का अर्थ $\{r, s\}$ और $\llbracket \text{ता है} \rrbracket$ का अर्थ शून्य (यानी $\{\varphi\}$) माना गया है। इस प्रकार, हम अपनी भाषाओं में सहायक-क्रियाओं को अर्थ हीन संरचना मान रहे हैं।

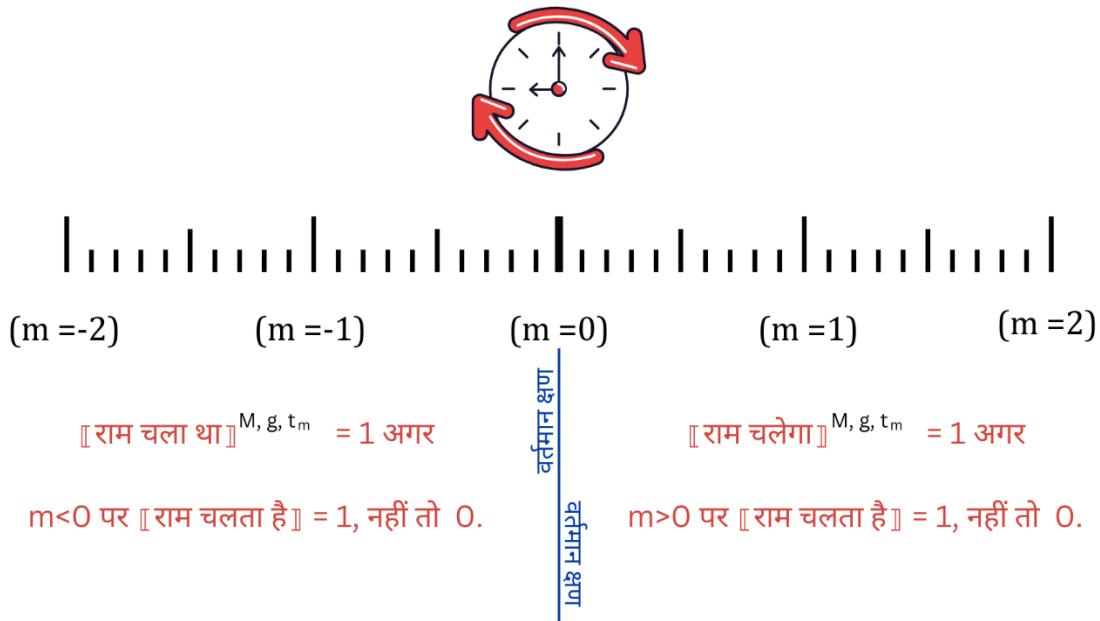


ऊपर जैसे कई मॉडल M हो सकते हैं जहाँ भाषाई इकाइयों को भिन्न-भिन्न अर्थ दिया जा सकता है। इन भाषाई इकाइयों पर हम व्याख्या-फ़ंक्शन (g) का उपयोग करते हैं जो मॉडल M के अंतर्गत काम करता है। प्रत्येक wff (यानी सिंटेक्स नियमों से बने हुए वाक्य) के अर्थ कई M से निकाले जा सकते हैं। यदि कल कोई M को बदल देगा तो वाक्य से निकलने वाले अर्थ और उनके आधार पर किये जाने वाले तर्क भी बदल जायेंगे।

यह प्रक्रिया वास्तविकता में "मानव के मस्तिस्क में शब्द के अर्थ को निकालने वाली प्रक्रिया" को कॉपी करने जैसा है। अब यदि कोई व्यक्ति किसी शब्द का कभी व्यवहारिक अर्थ लेता है या कभी व्यंग्यात्मक अर्थ लेता है, तो वो वास्तविकता में मॉडल M को ही बार-बार बदल रहा है।

- अर्थ निकालने के लिए संकल्पना ४**

अतीत और भविष्य काल-आधारित wff_s के अर्थ की गणना ऐखिक समय अनुक्रम का उपयोग करके की जाती है। यदि t_0 वर्तमान क्षण से मेल खाता है, जिसे 'ता हैं' सहायक द्वारा दर्शाया जाता है, तो t_0 के पहले के क्षण और t_0 के बाद के क्षण क्रमशः अतीत और भविष्य काल का प्रतिनिधित्व करते हैं।



४. मशीन और भाषा के तालमेल से भविष्य की कुछ मुख्य संभावनाएँ:

यह शोध मशीन को मुख्य रूप से गणितीय संभाव्यता (probability) से जुड़ी समस्याओं को व्याकरण नियमों (सिंटेक्स और सिमेंटिक्स) के सिद्धांतों से जोड़कर बेहतर उत्तर या बेहतर निर्णय लेने में सक्षम बनाएगा। इसका एक सबसे बड़ा उपयोग ए.आई (AI), एन.एल.पी (NLP), एम.एल. (ML), एल.एल.एम (LLM) आदि से बने उत्पादों में हो सकता है। तेजी से विकसित हो रही ये सभी तकनीकें उन कार्यों को स्वचालित कर सकती हैं जो पहले मनुष्यों द्वारा किए जाते थे। विशेष रूप से भाषा से जुड़े प्रयोग नीचे दी गई गढ़ तालिका में बताए गए हैं:

- **टेक्स्ट जनरेशन-** इस अनुप्रयोग में सामग्री निर्माण, रचनात्मक लेखन या यहाँ तक कि सॉफ्टवेयर के लिए कोड बनाने के लिए इस शोध का उपयोग किया जा सकता है।
- **ट्रांसलेशन-** इस अनुप्रयोग में टेक्स्ट या इमेज इनपुट का उपयोग करके टेक्स्ट को एक भाषा से दूसरी भाषा में अनुवाद किया जाता है।
- **टेक्स्ट सारांश-** इस अनुप्रयोग में लंबे लेखों, दस्तावेजों या वेब पेजों के सारांश तैयार करने के लिए NLP का उपयोग किया जाता है।
- **भावना विश्लेषण-** किसी कथन के पीछे की भावना को निर्धारित करने के लिए, चाहे वह सकारात्मक हो, नकारात्मक हो या तटस्थ (मुख्य रूप से सोशल मीडिया मॉनिटरिंग और ग्राहक प्रतिक्रिया विश्लेषण में हिंदी भाषा पे किया गया यह शोध उपयोग किया जाता है।
- **चैटबॉट और वर्चुअल असिस्टेंट-** AI द्वारा संचालित, और प्राकृतिक भाषा वार्तालाप करने के लिए कृत्रिम एवं सृजनात्मक भाषा का उपयोग किया जाता है।
- **टेक्स्ट वर्गीकरण-** टेक्स्ट को पूर्वनिर्धारित श्रेणियों या लेबल में बाँटने के लिए NLP का उपयोग किया जाता है (स्पैम का पता लगाने, सामग्री टैगिंग आदि जैसे कार्यों के लिए यह विशेषकर सहायक)।
- **भाषा स्थानीयकरण-** विभिन्न क्षेत्रों और संस्कृतियों के अनुरूप सामग्री का अनुवाद और अनुकूलन करने के लिए NLP का उपयोग किया जाता है, और इस प्रकार यह सुनिश्चित करता है कि संदेश सांस्कृतिक रूप से संवेदनशील हैं।
- **भाषण पहचान-** बोली गई भाषा को टेक्स्ट में बदलने के लिए, वॉयस असिस्टेंट और स्वचालित ट्रांसक्रिप्शन सेवाओं जैसे अनुप्रयोगों को सक्षम करने के लिए।
- **व्याकरण और शैली सुधार-** व्याकरण और शैली में भाषा-विशेष्ट सुधारों का सुझाव देकर उपयोगकर्ता के लेखन या बोलने के शैली को बेहतर बनाने के लिए।
- **संवाद और कथाएँ-** वीडियो गेम, सिमुलेशन और आभासी वातावरण के लिए संवाद और कथाओं का निर्माण।
- **भाषा संरक्षण-** अनुवाद, ट्रांसक्रिप्शन और दस्तावेज़ी सेवाओं के प्रावधान के माध्यम से लुप्तप्राय भाषाओं को संरक्षित करने में सहायता के लिए।
- **भाषा शिक्षण-** संबंधित भाषा के अनुरूप स्पष्टीकरण, उदाहरण और अभ्यास का उपयोग करके भाषा सीखने में सहायता प्रदान करना।

५. निष्कर्ष

यह शोध कार्य एक गणितीय मॉडल का उपयोग करके व्याकरण और शब्दार्थ विश्लेषण का कई उदाहरण प्रस्तुत करता है। प्रस्तावित व्याकरण और वाक्य-विन्यास के नियमों से बनाया गया हिंदी पार्सर शब्दिक और हिंदी के लिंग मानदंडों को सफलतापूर्वक स्पष्ट करता है - जैसे कि पुलिंग 'राम' के साथ 'चलती है' जैसी क्रियाओं को हमारा व्याकरण

सिंटेक्स के स्तर पे ही हटा देता है। विशेष रूप से शब्दक्रम जैसे कि हिंदी की SOV (Subject-Object-Verb) व्यवस्था का भी पूर्ण पालन दर्शाया गया है। भाषाई इकाइयों का सिंटेक्स-ट्री में "ऊपर से नीचे" की तरफ जाता हुआ व्यवस्थीकरण होता है। अतः हमने जो भी पार्सिंग नियम विकसित किए हैं, वे वाक्यों की स्व-व्यवस्था को संभालते हैं। यह क्षमता हिंदी वाक्यों के संरचनात्मक पहलुओं को समझने के लिए महत्वपूर्ण कदम है। हिंदी भाषा में प्रस्तावात्मक/कथन तर्क (propositional logic) और लिंग (gender) के नियमों को प्रभावी ढंग से हमारे मॉडल M में शामिल किया गया था। यह विश्लेषण हिंदी भाषा को औपचारिक और वैज्ञानिक दृष्टिकोण प्रदान करता है। 'अप्रतीकात्मक तर्क' (Asymbolic logic) जैसे कि मशीन लर्निंग (ML) और डीप लर्निंग (DL) को 'प्रतीकात्मक तर्क' (symbolic logic) के साथ मिलाने के लिए अपने आप में किया गया यह एक योगदान है। इस विषय पर अभी पूरे विश्व में वैज्ञानिक शोध चल रहा है और हिंदी की और से किया गया यह प्रयास पहला माना जा सकता है।

५. संदर्भ:

- Chomsky, Noam. "Syntactic structures." In *Syntactic Structures*. (2009). De Gruyter Mouton.
- Dowty, D. (1982). Grammatical relations and Montague grammar. In *The nature of syntactic representation* (pp. 79-130). Dordrecht: Springer Netherlands.
- Dowty, D. R., Wall, R. E., Peters, S., Dowty, D. R., Wall, R. E., & Peters, S. (1981). The Grammar of PTQ. *Introduction to Montague Semantics*, 179-251.
- Dowty, D. R., Wall, R., & Peters, S. (2012). *Introduction to Montague semantics* (Vol. 11). Springer Science & Business Media.
- Montague, R. (1975). English as a formal language. In J. Moravcsik (Ed.), *Logic and philosophy for linguists: A book of readings* (pp. 94-121). Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783111546216-007>
- Partee, B. (1984). Compositionality. *Varieties of formal semantics*, 3, 281-311.
- Partee, B. (1973). Some transformational extensions of Montague grammar. *Journal of Philosophical Logic*, 2, 509-534.