हिंदी वाक्यों में कर्ता अभिज्ञानक

(Subject Recognizer in Hindi Sentences)

महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय में एम.फिल. (कंप्यूटेशनल भाषाविज्ञान)

उपाधि के पाठ्यक्रम संबंधी आवश्यकता की आंशिक परिपूर्ति हेतु प्रस्तुत

लघु शोध-प्रबंध

शोधार्थी

दिब्य शंकर मिश्र

एम.फिल. (कम्प्यूटेशनल भाषाविज्ञान)

सत्र: 2012-13

शोध-निर्देशक

प्रो. विजय कुमार कौल

निदेशक

प्रौद्योगिकी अध्ययन केन्द्र एवं

कंप्यूटेशनल भाषाविज्ञान

सह शोध-निर्देशक

श्री जगदीप सिंह दांगी

एसोसिएट प्रोफेसर

कंप्यूटेशनल भाषाविज्ञान

कंप्यूटेशनल भाषाविज्ञान, भाषा विद्यापीठ

महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय

(संसद द्वारा पारित अधिनियम 1997 अंतर्गत, क्रमांक 3 के अंतर्गत स्थापित)

पोस्ट बॉक्स सं. -16, गांधी हिल्स, वर्धा-442005, (महाराष्ट्र)

फोन :07152-251661, फैक्स :07152-230903

ई:मेल-hindiuni_wda@sancharnet.in,info@hindivishwa.org

वेबसाइट :www.hindivishwa.org

विषयानुक्रम

पृ. सं

प्रस्तावना **2-5**

अध्याय-1 हिंदी वाक्य व्यवस्था : स्वरूप, परिभाषा एवं प्रकार **7-32**

1.1 वाक्य : स्वरूप एवं परिचय

1.2 वाक्य : परिभाषा

1.3 वाक्य : प्रकार

1.3.1 संरचनागत वर्गीकरण

(1) सरल वाक्य

(2) असरल वाक्य

(1.3.1.2.1) संयुक्त वाक्य

(1.3.1.2.2) मिश्र वाक्य

1.3.2 अर्थगत वर्गीकरण

अध्याय-2 हिंदी कारक व्यवस्था **34-48**

2.1 कारक : अर्थ एवं स्वरूप

2.2 कारक : परिभाषा

2.3 कारक : प्रकार

2.4 कर्ता कारक

अध्याय-3 हिंदी कर्ता अवस्थिति निर्धारण **50-81**

3.1 परसर्ग द्वारा

3.2 प्रकार्य द्वारा

3.3 संज्ञा पद द्वारा

अध्याय-4 कर्ता अभिज्ञानक का डाटा : संरचना एवं प्रबंधन 83-98

4.1 डाटा : परिभाषा एवं प्रक्रिया

4.2 डाटा : संरचना, प्रकृति एवं प्रबंधन

4.3 कर्ता अभिज्ञानक का डाटाबेस : डिजाइन

अध्याय-5 हिंदी वाक्यों में कर्ता अभिज्ञानक: संरचना एवं निर्माण 100-113

5.1 कर्ता अभिज्ञानक सॉफ्टवेयर आर्किटेक्चर

5.1.1 कर्ता अभिज्ञानक का प्रोग्रामिंग अवधारणा

5.1.2 कर्ता अभिज्ञानक प्रणाली संरचना

5.1.3 कलन विधि (Algorithm)

5.1.4 प्रवाह चित्र (Flow-Chart)

51.5 कर्ता अभिज्ञानक का कार्य प्रणाली

5.2 कर्ता अभिज्ञानक का अंतरापृष्ठ प्रस्तुतीकरण

5.2.1 डिजाइनिंग फार्म

5.2.2 रनिंग फार्म

5.2.3 कोडिंग

उपसंहार 115-117

संदर्भ सूची 119-121

संस्थापन निर्देश 122

कंप्यूटर की आवश्यकताएं 122

प्रस्तावना

मनुष्य भाषा के द्वारा अपने भाव, विचार और अनुभव दूसरों के सामने व्यक्त करता है और उनके भाव, विचार और अनुभव ग्रहण करता है। भाषा ही वह जरिया बनती है जिससे पारस्परिक संप्रेषण के संबंध स्थापित किए जा सकते हैं। संप्रेषण के यही संबंध हैं, जो मानव समाज को विकासक्रम में यहां तक लाए हैं। भाषा मुख से बोले जाने वाले शब्दों और वाक्यों का वह समूह है जिसके माध्यम से मन की बात कही या बताई जाती है। जब हम भाषा के जरिए किसी से कुछ कहते हैं तो उसके प्रत्युत्तर का साधन भी भाषा ही बनती है और एक संप्रेषणीयता कायम हो पाती है। दरअसल यह एक तरह से दो व्यक्तियों, समुदायों, समाजों के बीच किसी भी तरह के संबंध बनने का पहला माध्यम होती है। यह सब तब संभव हो पाता है जब हम भाषा की एक संरचना में जा कर वाक्यों का निर्माण करते हैं। वाक्य निर्माण के जितने भी तत्व हैं उन्हें जब उचित स्थानों पर स्थित किया जाता है तब यह संभव बन पाता है कि भाषा कोई अर्थवत्ता प्रदान करे। वाक्य भाषा में व्याकरणिक संरचना की महत्तम इकाई होता है, जिसके मुख्यतः दो अवयव या भाग होते हैं- उद्देश्य और विधेय। भाषा में वाक्यों के विभिन्न प्रकार हैं जिनका वर्णन इस अध्ययन में किया गया है, परंतु सरल वाक्य भाषा संप्रेषण के प्रथम चरण के तौर पर होते हैं। अध्ययन के तहत सरल वाक्य को ही आधार बनाया गया है। यदि वाक्यों के दो भागों पर हम चर्चा करें तो यह पाते हैं कि जिस वस्तु के विषय में कुछ कहा जाता है उसे सूचित करने वाले शब्दों को उद्देश्य कहते हैं तथा उद्देश्य के विषय में जो विधान किया जाता है उसे विधेय कहते हैं। उद्देश्य भारतीय पारंपरिक चिंतन में दी गई वह अवधारणा है जिसके द्वारा सामान्यतः कर्ता का प्रकार्य किया जाता है। अध्ययन के तहत कर्ता और प्रकार्य की विस्तृत चर्चा की गयी है। अध्ययन के दौरान हमने देखा कि भाषा के विभिन्न तत्वों को लेकर विद्वानों के अलग-अलग मत भी हैं। यह भी देखा कि विद्वानों ने वाक्य में निहित कर्ता को ही उद्देश्य माना है। सरल वाक्यों में इसके उदाहरण स्पष्ट तौर पर देखे जा सकते हैं। जब यह कहा या बोला जाता है कि घोड़ा दौड़ रहा है, लड़के ने किताब पढी, लड़का मार खाकर आ गया तो यहां घोड़ा, लड़का और लड़की उद्देश्य हैं और साथ ही ये कर्ता का भी कार्य कर रहे हैं। इस तरह से सरल वाक्य के ऐसे कई संदर्भ दिखते हैं जहां कर्ता और उद्देश्य एक ही हो जाते हैं। सरल वाक्य वाक्यों का एक प्रकार है जबकि इस अध्ययन के अंतर्गत हम देख सकते हैं कि वाक्यों की जो श्रेणियां बनाई गयी हैं, वह हिंदी में तीन तरह के वाक्य की रचना

करती हैं। जिनमें सरल वाक्य, संयुक्त वाक्य और मिश्र वाक्य के भी वर्णन किए गए हैं, परंतु अध्ययन में सरल वाक्य को ही आधार बनाया गया है।

सरल वाक्यों में कर्ता को पहचानने में आसानी होती है। क्योंकि इसमें एक उद्देश्य और एक ही विधेय होता है। वाक्य की संरचना भाषा के विकास के साथ जटिल होती गयी और वाक्य में कई ऐसे तत्व जुड़ते गए जिसने वाक्य की संरचना और तत्वों की स्थिति को और जटिल बनाया। जब हम तकनीक की दृष्टि से इस पर चर्चा करते हैं तो यह तकनीक प्रयोग के आधार पर कई जटिलताएं खड़ा करता है। मसलन वाक्य में किस तत्व की स्थिति क्या हो इसका निर्धारण सुनिश्चित ही करना होता है, क्योंकि तकनीक बगैर संरचना को जाने अक्षम हो जाती है। इस लिहाज से संप्रेषणीयता के विकास और सुगमता ने दूसरी तरफ विकसित हो रही तकनीक के सामने कई चुनौतियां खड़ी की। वह चुनौती यही थी कि भाषा के तहत बदलती शैली को कैसे संरचना बद्ध किया जाए। परंतु सरल वाक्यों के साथ यह तकनीकी सुगमता स्थापित करते हुए दिखती है क्योंकि यहां संरचना निर्धारित है और पदों व तत्वों की स्थितियां भी निर्धारित होती हैं। यहां वाक्य में इतने कम तत्व जुड़े होते हैं और संप्रेषणीयता इतनी स्पष्ट होती है कि इसे संरचनागत करने में कोई खास परेशानी नहीं दिखती। सरल वाक्यों में ज्यादातर वाक्यों का पहला शब्द कर्ता के रूप में होता है। जब हम सरल वाक्यों के तहत यह कहते हैं कि मोहन बुद्धिमान है, बिजली चमकती है, मोहन आम खा रहा है, तो यहां सब कुछ बेहद स्पष्ट है। कम तत्वों के साथ स्पष्ट तौर पर संप्रेषणीयता है। जहां उपर्युक्त वाक्यों में प्रथम शब्द कर्ता की भूमिका में है। उसको ही नियमबद्ध करने के लिए '**हिंदी वाक्यों में कर्ता अभिज्ञानक**' का विकास किया गया है, जो प्राकृतिक भाषा संसाधन की दृष्टि से महत्वपूर्ण है।

उद्देश्य- भाषा के साथ तकनीक को जोड़ने के लिए यह जरूरी है कि वाक्यों के प्रकार को समझा जाए। अध्ययन का उद्देश्य यह है कि यह वाक्यों की पहचान कर सरल वाक्यों में कर्ता के स्थान निर्धारण को चिन्हित करे। अध्ययन वाक्य के विभिन्न संदर्भों की चर्चा करते हुए इस उद्देश्य की पूर्ति करता है कि किस तरह से सरल वाक्य होते हैं और इनमे किन तत्वों के समावेश होने की संभावना बनती है। साथ ही सरल वाक्यों में किन तत्वों की स्थितियां क्या होंगी, इसकी भी पड़ताल की गयी है। इससे यह स्पष्ट होने की संभावना बनती है कि वाक्य में संरचना के विकास को भी समझा जा सकता है। प्रस्तुत लघु शोध-प्रबंध के कार्य का उद्देश्य हिंदी वाक्यों मे कर्ता

अभिज्ञानक का विकास करना है जिससे सरल वाक्यों में आए हुए कर्ता की पहचान की जा सके।

विषयवस्तु

अध्ययन के अंतर्गत जिन विषयों पर चर्चा की गयी है उसमें सरल वाक्य की व्यवस्था किस तरह से होती है? सरल वाक्य कैसे बनते हैं और इनका स्वरूप क्या होता है? इन पहलुओं पर चर्चा करते हुए अध्ययन के अंतर्गत उनके आसपास के ऐसे कई संदर्भों को रखा गया है जिसको समझे बिना या जिनकी सापेक्षता में ही दूसरे वाक्य को समझा जा सकता है। इसलिए सरल वाक्य की चर्चा करते हुए अन्य वाक्यों के प्रकारों पर भी इसके अंतर्गत चर्चा की गयी है। सरल वाक्य व अन्य प्रकार के वाक्यों को यहां परिभाषित किया गया है। वाक्यों के संदर्भ में विभिन्न विद्वानों ने जो विचार दिए हैं उनको रखा गया है ताकि यह समझा जा सके कि वे एक-दूसरे से किस तरह की मत भिन्नता रखते हैं। अध्ययन के अंतर्गत हिन्दी कारक व्यवस्था को भी शामिल किया गया है, जिसमे कर्ता, कारक और इनके विभिन्न पक्षों पर चर्चा की गयी है। साथ ही परसर्गों के प्रकार और परसर्गों की स्थिति व उनके प्रयोग को कुछ उदाहरणों से स्पष्ट करने और समझने का प्रयास किया गया है। संज्ञा पद की स्थिति पर भी विस्तृत चर्चा अध्ययन के तहत की गयी है।

शोध प्रविधि

प्रस्तुत लघु शोध-प्रबंध में विश्लेषणात्मक शोध पद्धति का प्रयोग किया गया है। इसके साथ ही भाषा वैज्ञानिक विश्लेषण संबंधित सिद्धांतों का प्रयोग करते हुए प्राकृतिक भाषा संसाधन से लेकर प्रोग्रामिंग भाषा C++ के माध्यम से नियम आधारित ‘हिंदी वाक्यों में कर्ता अभिज्ञानक’ प्रणाली का विकास किया गया हैं। इस पूरी विश्लेषणात्मक पद्धति में सोशल लिंग्विस्टिक के आधार पर भी चर्चा की गयी है।

अध्यायीकरण के लिहाज से इसे किस तरह रखा गया है यह स्पष्ट करना ज्यादा उचित लगता है। शोध कार्य में हिंदी वाक्यों में कर्ता अभिज्ञानक को केंद्र में रखा है। अध्याय योजना करते समय प्रथम अध्याय में हिंदी वाक्य व्यवस्था : स्वरूप, परिभाषा एवं प्रकार की चर्चा करते हुए विभिन्न क्षेत्रों की संक्षिप्त रूपरेखा प्रस्तुत करने का प्रयास किया गया है। दूसरे अध्याय में हिंदी कारक व्यवस्था के अंतर्गत अर्थ एवं स्वरूप, प्रकार एवं कर्ता कारक के बारे में संक्षिप्त रूपरेखा प्रस्तुत किया गया है। तीसरे अध्याय में हिंदी कर्ता अवस्थिति निर्धारण के अंतर्गत परसर्ग, प्रकार्य (function) एवं संज्ञापद के माध्यम से कर्ता की अवस्थिति निर्धारण किया गया है।

चौथे अध्याय में तकनीकी पक्ष के डाटाबेस : निर्माण एवं प्रबंधन की उपयोगिता और सूचनाओं को संचित करने की विधि का उल्लेख किया गया है।

पांचवे अध्याय में कर्ता अभिज्ञानक प्रणाली : संरचना एवं निर्माण प्रक्रिया के अंतर्गत सॉफ्टवेयर संरचना, प्रोग्रामिंग अवधारणा, कलन विधि (Algorithm), प्रवाह चित्र (Flow-Chart),निर्माण प्रक्रिया, प्रणाली का अंतरापृष्ठ चित्र एवं प्रणाली मूल्यांकन आदि का उल्लेख किया गया है। साथ में यह प्रोग्राम कैसे कार्य करता है उसके प्रारूप को भी दर्शाया गया है।

इन अध्यायों के उपरांत उपसंहार को लिखा गया है, तत्पश्चात 'संदर्भ ग्रंथ सूची' में संकलित एवं अध्ययित की गई पुस्तकों की सारणी, प्रणाली से संबंधित संस्थापन नियम, प्रणाली संस्थापन की आवश्यकता इत्यादि का उल्लेख किया गया है।

प्रस्तुत लघु शोध प्रबंध में अतंरराष्ट्रीय अंकों का प्रयोग किया गया है तथा पाद टिप्पणी प्रत्येक पृष्ठ पर दी गई है।