मानक हिंदी के लिए संज्ञा व्युत्पादक रूपिम जनरेटर

“(Derivational Morphological Generator Noun in Standard Hindi)”

महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय , वर्धा में एम.फिल. कंप्यूटेशनल भाषाविज्ञान की उपाधि हेतु प्रस्तुत

लघु शोध-प्रबंध

शोधार्थी

धीरेन्द्र यादव

नामांकन संख्या:

2014/01/202/002

सत्र: 2014-15

कंप्यूटेशनल भाषाविज्ञान विभाग, भाषा विद्यापीठ

महात्मा गांधी अंतरराष्ट्रीय हिंदी विश्वविद्यालय

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

गांधी हिल्स, वर्धा (महाराष्ट्र) 442005

शोध-सारांश

आज का दौर सूचना प्रौद्योगिकी का है और इस दौर में प्रत्येक व्यक्ति अपनी निज भाषा में इसका उपयोग करना चाहता है। इसके लिए प्राकृतिक भाषा संसाधन तकनीक के द्वारा कई सॉफ्टवेयर उपकरण विभिन्न क्षेत्रिय भाषाओं हेतु विकसित हो रहें हैं। विश्व में तेजी से हो रहे संगणकीय भाषा के विकास ने प्रत्येक भाषा की विशेषता को बारीकी से समझने पर जोर दिया है। हिंदी भाषा के माध्यम से आज भारत के प्रत्येक गाँव को एक साथ जोड़ने के लिए वैश्विक ग्राम की संकल्पना को साकार करने में सरकार भी प्रयत्नशील है। हिंदी आज दुनिया की प्रमुख भाषाओं में से एक है जिसे अहिंदीतर भाषी भी सीखने में रुचि दिखा रहे हैं।

प्रस्तुत शोध कार्य में हिंदी भाषा के संज्ञा रूपों का विश्लेषण/संश्लेषण कर **"मानक हिंदी के लिए संज्ञा व्युत्पादक रूपिम जनरेटर (Derivational Morphological Generator Noun in Standard Hindi)"** का निर्माण किया गया है। जैसा कि हम जानते हैं व्युत्पादक रूपिम जिसमें प्रत्यय या उपसर्ग लगाने पर उस शब्द के अर्थ में परिवर्तन होता है एवं उस शब्द की व्याकरणिक कोटि भी बदल जाती है। प्रस्तुत शोध कार्य के अंतर्गत हिंदी भाषा के संज्ञा शब्दों का संकलन करतथा उनमें लगने वाले विभिन्न प्रत्ययों के आधार पर विश्लेषण एवं संश्लेषण किया गया है। चूंकि किसी भी संज्ञा के धातु अथवा प्रातिपदिक शब्दों में प्रत्यय के योग से उसकी व्याकरणिक कोटि का बदल जाना एक स्वाभाविक प्रक्रिया है। अत: प्रस्तुत शोध-प्रबंध में संगणकीय मॉडल का निर्माण कर विभिन्न नियम बनाकर संज्ञा व्युत्पादक रूपिम जनरेटर हेतु एक सॉफ्टवेयर प्रणाली का निर्माण किया गया है, जो संज्ञा के मूल शब्द में प्रत्यय जोड़कर उसका विश्लेषण और संश्लेषण कर उसकी व्याकरणिक कोटियों को दिखाता है। यह प्राकृतिक भाषा संसाधन का एक ऐसा अनुप्रयुक्त क्षेत्र है जोहिंदी भाषा के कम-से-कम धातु/प्रातिपदिक से प्रत्यय के योग से ज्यादा-से-ज्यादा शब्दों का व्युत्पादन करता है। इसका उपयोग वर्तनी जांचक, POS टैगर, मशीनी अनुवाद तथा सूचना पुनरप्राप्ति आदि में किया जा सकता है।

बीजशब्द (Keywords):- व्युत्पादक रूपिम, हिंदी भाषा, सॉफ्टवेयर प्रणाली, जनरेटर, संज्ञा, प्रत्यय

Summary of Research

Present time is the time of information technology, in this period all the human being wants to use the same i.e. IT in his daily curricular activities. By implementation of NLP techniques the regional Languages related software and tools are being developing. The fast development of computational Linguistics focused on serious studies of all languages in the world. The Government is also trying to make to digital India for connecting villages on one platform like a global village with the help of Hindi Language. Hindi language is one of the most famous languages in the world, now day's non Hindi speakers are also interested to learning the same.

This is an applied area of NLP. In this process, with the addition of suffixes/prefixes to the root word in Hindi that might be expands or increased in new form which is either inflectional or derivational. Unfortunately, there is limited work has been done on derivational morpheme, analysis and synthesis of noun form. In the proposed researched **"Derivational Morphological Generator Noun in standard Hindi"** has been done. As we know in derived morpheme after the addition of suffixes or prefixes in the root word the meaning or its GC (grammatical category) will be changed. In this research, having the collection of nouns & suffixes which can add in the root has been analyzed and rules for the same derivational morpheme have been made changing of meaning is very common process after addition of the prefix/suffix in any root form. In this dissertation/research, on the basis of computational model of Derivational Morphological Generator a tool has been developed which shows the grammatical category of new generated word i.e. derived noun. This research or research-tool is based on Hindi so it may be very helpful in various Hindi tools such as spell checker, Machine Translation, POS Tagger, information retrieval etc.

Keywords:-

Derivational Morpheme, Hindi, Language, Software, Generator, Noun, Prefix/Suffix