हिन्दी शब्दतन्त्र

संगणक विज्ञान एवं अभियांत्रिकी विभाग
भा. प्रौ. सं. मुम्बई
(http://www.cfilt.iitb.ac.in/wordnet/webhwn/wn.php)

- प्रभाकर पाण्डेय, लक्ष्मी कश्यप, पुष्पक भट्टाचार्य

हिन्दी शब्दों के मध्य विभिन्न प्रकार के संबंधों को दर्शाने का एक संगणकीय शाब्दिक तन्त्र है - हिन्दी शब्दतन्त्र। यह पारम्परिक शब्दकोशों से अलग है। यह वैज्ञानिकता के आधार पर शब्दों के बारे में यथार्थ जानकारी उपलब्ध कराता है। इसकी संरचना अंग्रेजी शब्दतंत्र (English WordNet) पर आधारित है पर यह हिन्दी की विशिष्टता को समाहित किए हुए है।

हिन्दी शब्दतन्त्र में अर्थ की समानता के आधार पर पर्याय-समूह (Synset) का निर्माण एक स्पष्ट व्याख्या और वाक्य-प्रयोग के साथ किया जाता है ताकि शब्दों की अनेकार्थता के कारण अर्थ की स्पष्टता बाधित न हो। वास्तव में, ये पर्याय-समूह (Synset) ही हिन्दी शब्दतन्त्र के आधार हैं। फिलहाल अभी तक हिन्दी शब्दतन्त्र में संज्ञा, विशेषण, क्रिया एवं क्रिया-विशेषण वर्ग के शब्दों को ही स्थान दिया गया है।

आइए, हिन्दी शब्दतन्त्र की संरचना में योगदान देनेवाले प्रत्येक तत्वों से परिचित होते हैं :-

१. **पर्याय-समूह** (Synset) – इसमें किसी संकल्पना की यथार्थता को सूचित करनेवाले शब्दों को क्रम से बारम्बारता के

आधार पर व्याख्या और उदाहरण के साथ दर्शाया जाता है।

जैसे- गाय, गऊ, धेनु- सींगवाला एक शाकाहारी मादा चौपाया "गाय अपने बछड़े को दूध पिला रही है।"

२. **सत्ता-मीमांसा** (Ontology) – हिन्दी शब्दतन्त्र में शब्द-भेद के आधार पर शब्द की संकल्पना को सुस्पष्ट करने के लिए सत्ता-मीमांसा भी दी जाती है।

जैसे- गाय, गऊ, धेनु

- ⇨ पालतू पशु
- ⇨ पशु
- ⇨ सजीव
- ⇨ संज्ञा

हिन्दी शब्दतन्त्र में संबंधों की रूप-रेखा :-

हिन्दी शब्दतन्त्र संकल्पना पर आधारित है एवं ये संकल्पनाएँ भाव की दृष्टि से प्रत्यक्ष या अप्रत्यक्ष रूप से एक दूसरे से जुड़ी होती हैं। अस्तु इन संकल्पनाओं को एक दूसरे से जोड़ने के लिए अधिवाची, अधोवाची, अंगवाची, अंगीवाची, विपर्यायवाची आदि संबंधसूचक अवधारणाओं का उपयोग किया जाता है। इनमें से कुछ शब्दों के मध्य के संबंधों को तो कुछ पर्याय-समूहों के मध्य के संबंधों को दर्शाती हैं।

आइए इन संबंधसूचक अवधारणाओं को जानें –

१. **अधिवाची** (Hypernymy) **एवं अधःवाची** (Hyponymy) :-

पर्याय-समूहों के मध्य महत्वपूर्ण संबंध को दर्शानेवाली ये दोनों अवधारणाएँ एक दूसरे की पूरक हैं । यदि 'क' एक प्रकार का 'ख' है तो 'ख' 'क' का अधिवाचक और 'क' 'ख' का अधःवाचक है और ये संबंध अधिवाची और अधःवाची कहलाते हैं।

जैसे– गाय (क) ----- चौपाया (ख)

२. **अंगवाची** (Meronymy) **एवं अंगीवाची** (Holonymy) :-
ये संबंधसूचक अवधारणाएँ भी पर्याय-समूहों के मध्य होती हैं । यदि 'क' 'ख' का भाग है तो 'क' 'ख' का अंगवाचक और 'ख' 'क' का अंगीवाचक है और ये सम्बन्ध अंगवाची एवं अंगीवाची कहलाते हैं ।

जैसे– गाय (ख) ----- थन (क)

३. **विपर्यायवाची (**Antonymy**)** :-
यह संबंधसूचक अवधारणा पर्याय-समूहों के मध्य न होकर शब्दों के मध्य होती है । यह अवधारणा दो शब्दों के मध्य अवस्था, कार्य, समय, गुण आदि के आधार पर विपरीत अर्थ को दर्शाती है ।

जैसे– बेटा – बेटी (लिंग)
पुत्र – पुत्री (लिंग)
दिन – रात (समय)

४. **श्रेणीकरण (**Gradation**)** :-
यह संबंधसूचक अवधारणा दो विपर्याय शब्द संकल्पनाओं के बीच अवस्था, कार्य, समय, गुण आदि के आधार पर एक तीसरी शब्द संकल्पना को दर्शाती है ।

जैसे– सुबह – दुपहर – शाम (समय)
लड़कपन – जवानी – बुढ़ापा (अवस्था)

५. **अपरिहार्यतावाची (**Entailment**)** :-
यह संबंध दो क्रिया पर्याय-समूहों के मध्य होता है । यदि 'क' क्रिया 'ख' क्रिया में निहित है तो 'क' 'ख' का अपरिहार्यतावाचक है और यह संबंध अपरिहार्यतावाची कहलाता है ।

जैसे– 'खर्राटा लेना' क्रिया में 'सोना' क्रिया निहित है ।

६. **प्रकारवाची (**Troponymy**)** :-
यह संबंध भी केवल दो क्रिया पर्याय-समूहों के मध्य होता है । इसमें एक क्रिया किसी दूसरी क्रिया के किसी विशेष ढंग को दर्शाती है ।

जैसे– 'मुस्कुराना' 'हँसना' का प्रकारवाचक है और यह संबंध प्रकारवाची कहलाता है ।

७. **प्रेरणार्थक क्रिया (**Causative Verb**)** :-
इसमें मूल एवं प्रेरणार्थक क्रियाओं के मध्य संबंध दर्शाया जाता है ।

जैसे– चलना – चलाना
पढ़ना – पढ़ाना

शब्द-भेदों के मध्य संबंध :-

हिन्दी शब्दतन्त्र में कुछ संबंध शब्द-भेदों के मध्य भी दिए गए हैं। आइए, उनपर नजर डालें।

(क) **संज्ञा-पद एवं क्रिया-पद के मध्य संबंध** :-

१. *आन्तर-योग्यता निर्देशी क्रिया (Ability Link)* :-

इस संबंध के अन्तर्गत कोई क्रिया-पद किसी संज्ञा-पद के प्राकृतिक (मूल) गुण को दर्शाता है।

जैसे – मछली, मीन, मत्स्य (संज्ञा-पद)

⇨ तैरना, पैरना (क्रिया-पद)

२. *बाह्य-योग्यता निर्देशी क्रिया (Capability Link)* :-

इस संबंध के अन्तर्गत कोई क्रिया-पद किसी संज्ञा-पद के बाह्य या अर्जित योग्यता को दर्शाता है।

जैसे – व्यक्ति, मानस (संज्ञा-पद)

⇨ तैरना, पैरना (क्रिया-पद)

३. *कर्म निर्देशी क्रिया (Function Link)* :-

इस संबंध के अन्तर्गत कोई क्रिया-पद किसी संज्ञा-पद के कर्म को निर्दिष्ट करता है।

जैसे – अध्यापक, शिक्षक (संज्ञा-पद)

⇨ पढ़ाना, शिक्षा देना (क्रिया-पद)

(ख) **संज्ञा-पद एवं विशेषण-पद के मध्य सम्बन्ध** :-

१. *गुणवाची (Attribute)* :-

इस संबंध के अन्तर्गत कोई विशेषण-पद किसी विशेष संज्ञा-पद के गुणधर्म को सूचित करता है।

जैसे – बाघ (संज्ञा-पद) – मांसाहारी (विशेषण-पद)

२. *अर्थ-संकुचन संज्ञा (Modified Noun)* :-

कुछ विशेषण-पद केवल कुछ विशेष संज्ञा-पद या संज्ञा-पदों को ही विशेषित करते हैं। ऐसे विशेषण-पद एवं संज्ञा-पद, अर्थ संकुचन-संज्ञा सम्बन्ध द्वारा एक दूसरे से जुड़े होते हैं।

जैसे – पंखदार, पाँखदार (विशेषण-पद)

⇨ पक्षी, चिड़िया (संज्ञा-पद)

(ग) **क्रिया-पद एवं क्रियाविशेषण-पद के मध्य संबंध** :-

१. *अर्थ-संकुचन क्रिया (Modified Verb)* :-

कुछ क्रियाविशेषण-पद केवल कुछ विशेष क्रिया-पद या क्रिया-पदों को ही विशेषित करते हैं। ऐसे क्रियाविशेषण-पद एवं क्रिया-पद, अर्थ संकुचन-क्रिया सम्बन्ध द्वारा एक दूसरे से जुड़े होते हैं।

जैसे – रिमझिम-रिमझिम, (क्रियाविशेषण-पद)

⇨ बरसना, वर्षा होना (संज्ञा-पद)

२. *<u>से व्युत्पन्न (Derived from)</u>* :-

इस संबंधबोधक अवधारणा द्वारा यह दर्शाया जाता है कि व्युत्पन्न शब्द की व्युत्पत्ति किस मूल शब्द से हुई है।

जैसे – क्रमशः, क्रमानुसार, क्रमवार, सिलसिलेवार

⇨ क्रम, सिलसिला

इसप्रकार, हम देख सकते हैं कि शब्द की अर्थ-स्पष्टता और विषय-वस्तु के सटीक वर्गीकरण के साथ-साथ शब्दों के मध्य के आपसी संबंधों को दर्शानेवाला हिन्दी शब्दतन्त्र भाषा और संगणक का एक चमत्कारी सुमेल है जो शब्दों का उचित विश्लेषण करके प्रयोगकर्ताओं के आगे अथाह शब्द-व्यंजनों को परोसता है।

लक्ष्य :-

१.सभी हिन्दी शब्दों को शामिल करना।
२.संबंध दर्शाना।
३.मुहावरे, कहावतों आदि को भी शामिल करना।
४.शब्दों के लिंग सूचित करना।

चुनौतियाँ :-

(क) लिखित विषय-वस्तु

1. अच्छे शब्दकोशों एवं कारपोरा की कमी।
2. शब्दों के लाक्षणिक अर्थ। कहाँ तक ?

उदाहरण-

पाकिस्तान और हिन्दुस्तान के संबंधों में गरमाहट आ गई है।

3. समान शब्द के लिए विभिन्न शब्दकोशों में अलग-अलग संकल्पना-

उदाहरण- हिन्दी शब्द 'तंति' का अर्थ नालंदा हिन्दी शब्दकोश में 'गाय' दिया है जबकि आक्सफोर्ड हिन्दी अंग्रेजी शब्दकोश में रस्सी और बुनकर दिया है।

(ख) शब्दकोश में दिए गए शब्द-भेद और वाक्य-प्रयोग में अंतर-

उदाहरण **- उकड़ू**

शब्दकोश (विशेषण)

वाक्य-प्रयोग- 'वह उकड़ू बैठा है' (क्रिया विशेषण)

●शब्दकोश में एक शब्द के सभी अर्थों की कमी

उदाहरण- दोगुना, तीनगुना, चारगुना इत्यादि (क्रिया विशेषण)

समासिक शब्द:

●अधिक उत्पादनशीलता

उदाहरण -जीवनकला,जीवनक्रम,जीवनदृष्टि,जीवननिर्वाह,जीवनपध्दति,जीवनपुष्प,जीवनमान, जीवनमूल्य,जीवनवृक्ष,जीवनसूत्र,जीवनावश्यक ----**कहाँ तक लें ?**

●संयुक्त एवं सामासिक क्रियाएँ:- सीमा निर्धारण ?

उदाहरण- आरम्भ करना, छलांग मारना ।

<u>विशेष</u> :-

हिन्दी शब्दतन्त्र का निर्माण-कार्य अभी भी 'भारतीय प्रौद्योगिकी संस्थान मुम्बई' के 'संगणक विज्ञान एवं अभियांत्रिकी विभाग' में प्रा0 पुष्पक भट्टाचार्य (http://www.cse.iitb.ac.in/~pb/) के मार्गदर्शन में अविराम चल रहा है । अभी इसमें २६००० (छब्बीस हजार) से ऊपर पर्याय-समूह (Synset) हैं जो ५८००० (अट्ठावन हजार) यूनिक (एकमेव) शब्दों को समाहित किए हुए हैं । हिन्दी शब्दतन्त्र का आनलाइन उपयोग करने के लिए या इसे डाउनलोड करने के लिए कृपया निम्न कड़ी पर जाएँ – http://www.cfilt.iitb.ac.in/wordnet/webhwn/wn.php