

आर्ट्स अँड कॉमर्स कॉलेज, आष्टा

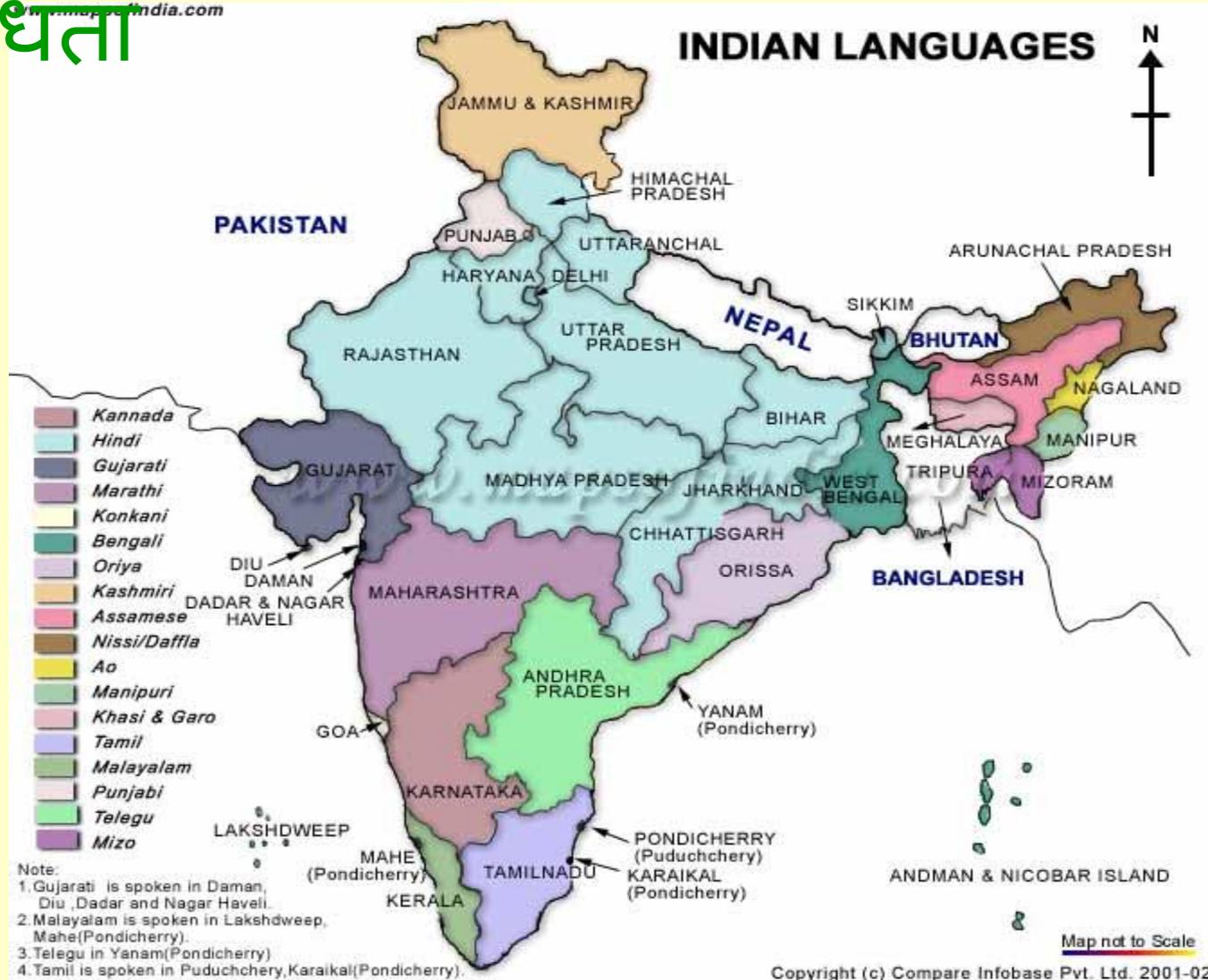
अध्यापक का नाम- विश्वास निवृत्ती पाटील

बी. ए. भाग-3

भाषा विज्ञान

भारत में महत् भाषा-विविधता

www.mapsofindia.com
धता



भाषा एवं भाषियों की संख्या

भाषा	जनसंख्या (2001 की जनगणना के आधार पर; लगभग)
हिंदी	450, 000, 000
मराठी	72, 000, 000
कोंकणी	7, 000, 000
संस्कृत	6000
नेपाली	13, 000, 000

भाषा एवं भाषियों की संख्या (क्रमशः)

भाषा	जनसंख्या (2001 की जनगणना के आधार पर; लगभग)
कश्मीरी	5, 000, 000
आसामी	13, 000, 000
तमिल	60, 000, 000
मलयालम	33, 000, 000
बोड़ो	1, 000, 000
मणिपुरी	1, 000, 000

प्रमुख भाषा प्रक्रिया पहल

- मुख्यतः सरकार की ओर से : सूचना एवं प्रौद्योगिकी मंत्रालय, मानव ससाधन एवं विकास मंत्रालय, विज्ञान एवं तकनीकी विभाग
- हाल में औद्योगिक संस्थानों द्वारा अत्यधिक रुझान : भारतीय भाषाओं पर केंद्रित प्रा.भा.प्र.(एनएलपी) का प्रयास
 - गूगल
 - माइक्रोसाफ्ट
 - आईबीएम शोध प्रयोगशाला
 - याहू
 - टीसीएस

प्राकृतिक भाषा प्रक्रिया (एन.एल.पी.), भा.प्रौ.सं. मुंबई शोध-दल,
सरकार और औद्योगिक संस्थानों द्वारा अत्यधिक समर्थित

हिन्दी शब्दतन्त्र

- शब्दतंत्र – एक शाब्दिक आँकड़ासंचय
- हिंदी शब्दतंत्र – अंग्रेजी शब्दतंत्र द्वारा प्रेरित
- संकल्पना के आधार पर निर्मित
- पर्याय-समूह – शब्दतन्त्र के मुख्य आधार
- विभिन्न वाक्यगत वर्गों के लिए विभिन्न व्यवस्था सिद्धांत

हिन्दी शब्दतन्त्र की संरचना

1. पर्याय-समूह

गाय, गऊ, गैया, धेनु

व्याख्या (उदाहरण सहित)

सींगवाला एक शाकाहारी मादा चौपाया

हिन्दू लोग गाय को गो माता कहते हैं ।

2. शब्द-भेद निर्धारण-संज्ञा, क्रिया, विशेषण, क्रिया-विशेषण ।

3. सत्ता-मीमांसा निर्धारण

पर्याय-समूह निर्माण संबंधी सिद्धांत

- **संक्षिप्तता**

- कम से कम पर्याय-शब्दों के प्रयोग से संकल्पना की एकार्थता का स्पष्टीकरण

- **व्यापकता**

- संकल्पना की एकार्थता को दर्शानेवाले सभी पर्याय-शब्दों को क्रम से उनकी बारम्बारता के आधार पर पर्यायवाची-समूह में प्रविष्टि

- **प्रतिस्थापनीयता**

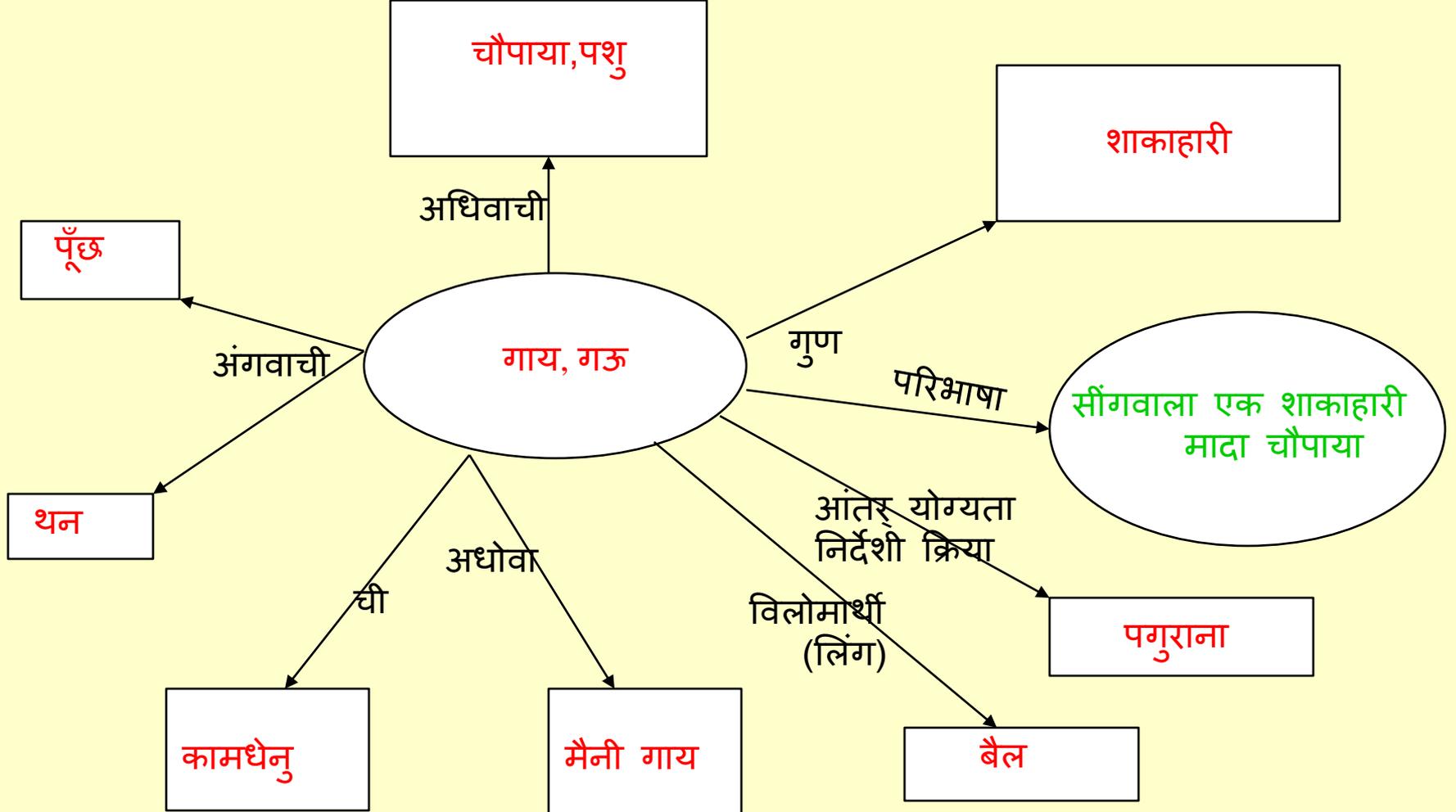
- पर्याय-समूह में आए हुए अधिकतर शब्दों का उदाहरण-वाक्य में बिना अर्थ परिवर्तन के प्रतिस्थापना संभव

- घर, गृह, मकान, सदन, शाला, आलय, धाम, निकेतन, वास्तु, पण-मनुष्यों का छाया हुआ वह स्थान, जो दीवारों से घेरकर बनाया जाता है "इस घर में पाँच कमरे हैं"

शब्दतंत्र में संबंध

- पर्यायवाची (Synonymy)
- अधिवाची (Hypernymy) / अधोवाची (Hyponymy)
- विलोमार्थी (Antonymy)
- अंगवाची (Meronymy) / अंगीवाची (Holonymy)
- श्रेणीकरण (Gradation)
- अपरिहार्यतावाची (Entailment)
- प्रकारवाची (Troponymy)

संज्ञा-रेखाचित्र



आँकड़े

पर्याय-समूह	33500
एकमेव शब्द	80400
संबंधयुक्त पर्याय-समूह	33500
हिंदी-अंग्रेजी योजित पर्याय-समूह	13000
खोजे गए शब्द	260000

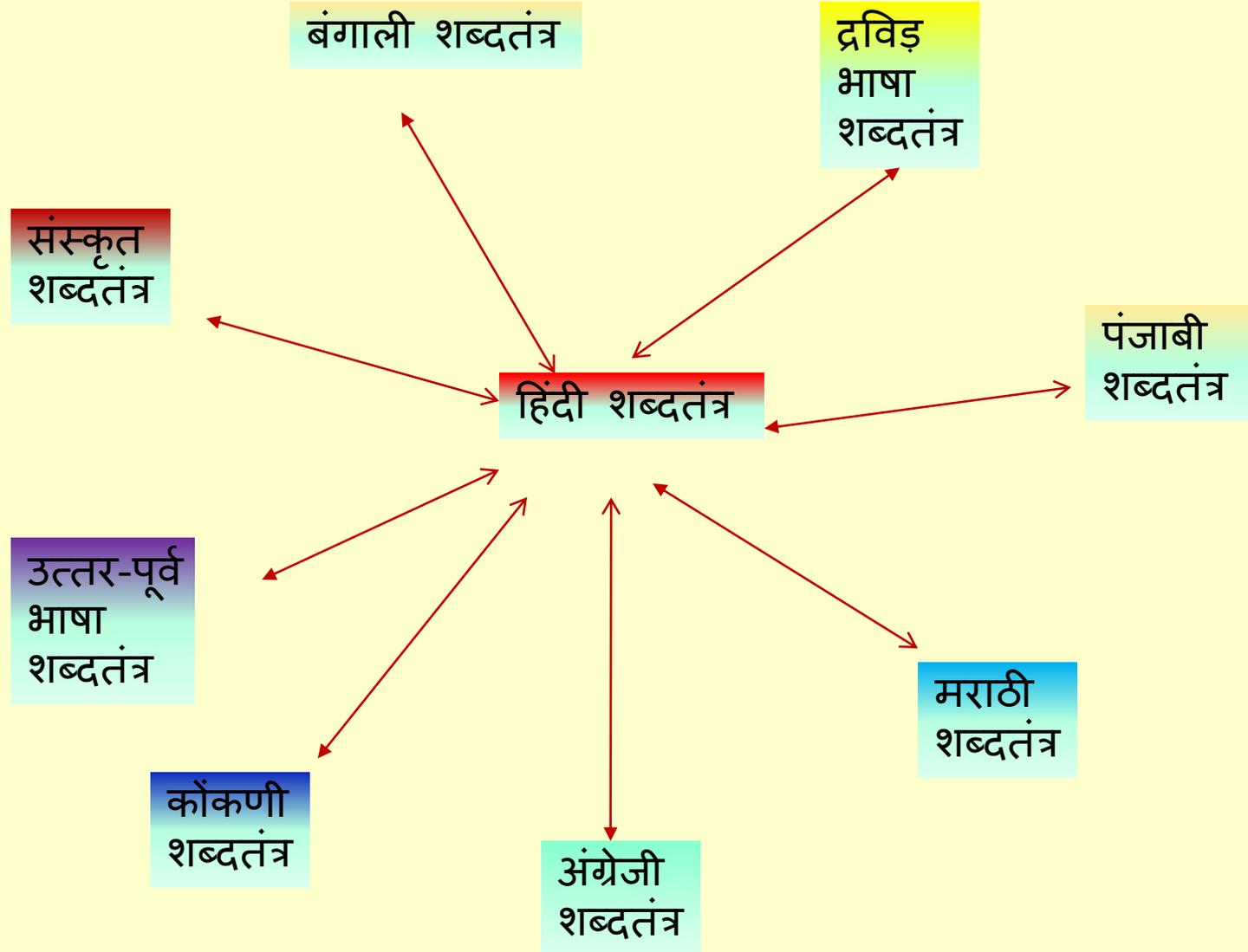
प्रभाव, उपयोगिता एवं प्रत्यक्षता

- उपयोग के लिए आनलाइन उपलब्ध (निःशुल्क)
- शोध कार्य हेतु निःशुल्क डाउनलोड की सुविधा
- भाषाविज्ञान आँकड़ा संघ (LDC), पेंसिलवेनिया विश्वविद्यालय, अमरीका द्वारा सार्वभौमिक उपलब्धता
- भारतीय भाषा खोज इंजिन पर काम करने के लिए गूगल द्वारा क्रीत व्यावसायिक अनुज्ञापत्र
- यूरोपीय भाषा संसाधन संगठन (ELRA) द्वारा जल्द उपलब्ध: यूरोप का भाषा आँकड़ा कोश
- भारतीय भाषाविज्ञान आँकड़ा संघ द्वारा उपलब्ध

प्रभाव, उपयोगिता एवं प्रत्यक्षता (क्रमशः)

- विश्वभर से दैनिक उल्लेख
- सन् 2006 से 2.5 लाख से अधिक शब्द-खोज
- 4600 से अधिक डाउनलोड
- कई भारतीय भाषाओं के शब्दतंत्र का केंद्रबिंदु
- भारतीय भाषा अनुवाद, संक्षेपण, भाषाओं के मध्य खोज के लिए कई शोधकर्ताओं द्वारा आधारभूत संसाधन के रूप में प्रयुक्त

अन्य भारतीय भाषाओं के शब्दतंत्रों की धुरी



संलग्न शब्दतंत्र

- अपरिमित शब्द संसाधन
- यांत्रिक अनुवाद, भाषा के मध्य खोज के लिए बहुत ही लाभकारी
- भाषा-अध्यापन, प्रशिक्षण, तुलनात्मक भाषाविज्ञान के लिए बहुत उपयोगी
- युरोशब्दतंत्र के समान परंतु विशिष्ट भारतीय भाषा अभिलाक्षणिकता के कारण समीक्षात्मक अंतर

शब्दतंत्र पर आधारित मानक अखिल भारती य शब्दकोश

शब्दार्थ	हिंदी	मराठी	बंगाली	उड़िया	तमिल
($W_1, W_2, W_3, W_4, W_5, W_6$)	($W_1, W_2, W_3, W_4, W_5, W_6$)	(W_1, W_2, W_3)	(W_1, W_2, W_3)	(W_1, W_2, W_3, W_4)	(W_1, W_2, W_3)
(sun)	(सूर्य, सूरज, भानु, भास्कर, प्रभाकर, दिनकर, अंशुमान, अंशुमाली)	(सूर्य, भानु, दिवाकर, भास्कर, रवि, दिनेश, दिनमणी)
(cub, lad, laddie, sonny, sonny boy)	(लड़का, बालक, बच्चा, छोकड़ा, छोरा)	(मुलगा, पोरगा, पोर, पोरगे)
(son, boy)	(पुत्र, बेटा, लड़का, लाल, सुत, बच्चा, सूत, नंदन, नन्दन, पूत, तनय)	(मुलगा, पुत्र, लेक, चिरंजीव, तनय)

सम्मान

- पी.के. पटवर्धन प्रौद्योगिकी विकास पुरस्कार, 2008
- हिंदी शब्दतंत्र पर आधारित बहुभाषीय आँकड़ासंचय निर्माण के लिए माइक्रो साफ्ट शोध भारत के द्वारा शोध अनुदान
- हिंदी शब्दतंत्र का घटक के रूप में अव्यवस्थित सूचना प्रबंधन के लिए आईबीएम भारत शोध अनुदान

अंतर्राष्ट्रीय सार्वभौमिक शब्दतंत्र सम्मेलन, 31 जनवरी से 4 फरवरी,

हिंदी शब्दतंत्र की सफलता के कारण भा. प्रौ.सं. मुंबई को अनुदत्त एक प्रमुख अंतर्राष्ट्रीय कार्यक्रम

Global Wordnet Conference 2010 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.cfilt.iitb.ac.in/~gwc2010/index.php

Most Visited Gmail - Summary: SM... Getting Started Calendar Dates Frame... Latest Headlines Gmail - [Fwd: TAM dic...

Cricinfo.com | Cricket ne... Gmail - Compose Mail - ... Global Wordnet Confer... Gmail: Email from Google UNL Explorer | UNL Intelli...

Final Call For Papers Paper Submission Important Dates Registration Accommodation Attendees

Global Wordnet Conference 2010

Navigate

- Home
- Final Call For Papers
- Paper Submission
- Important Dates
- Registration
- Accommodation
- Attendees
- Contact Us

Other Links

- Travel and Stay Guide
- Visa Information
- Committees
- Sponsorship
- Partners

Announcements

Submission deadline extended up to 5th October, 11.59 PM Pacific time. The site will close after that time.

