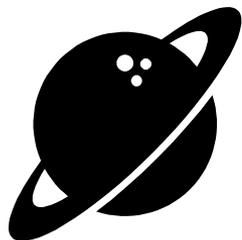
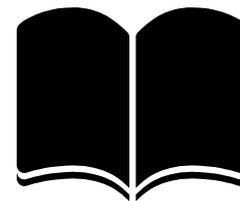


Hindi Wordnet for Language Teaching

- Experiences and Lessons Learnt



Hanumant Redkar¹, Rajita Shukla², Sandhya Singh¹,
Jaya Saraswati¹, Laxmi Kashyap¹, Diptesh Kanojia¹,
Preethi Jyothi¹, Malhar Kulkarni¹, Pushpak Bhattacharyya¹

¹Indian Institute of Technology Bombay (IITB), India

²Bennett University, Greater Noida, India



Tata Centre
Technology and Design, IIT BOMBAY



Outline

- Overview
- Motivation
- Language learning through association capturing
- A **Digital Aid** for language teaching and learning
- Process, Experiences and Lessons learnt
- Field test and User feedback
- Conclusion and Future work

Overview

- This paper reports the work related to making Hindi Wordnet available as a digital resource for language learning and teaching
- It presents the experiences and lessons that were learnt during the process
- Digital Aid, Hindi Shabdmitra presented here is based on modern pedagogical axioms and is aligned to the learning objectives of the syllabi of the school education in India
- It is a comprehensive language tool having audiovisual and grammatical features with multi-layered, multi-sensory information delivery

Motivation

- Hindi wordnet [1] is a large digital lexical database of a language in which information is organized around cognitive synonym sets or synsets
 - The underlying basis of such organization are the word association studies in psycholinguistics. Thus, subjects respond quicker than normal to the word 'nurse' if it follows a highly associated word such as 'doctor' [2]
 - Conventional sources of vocabulary learning (*viz.*, dictionaries , thesauri) do not have these relations due to the very nature of their composition
- The second motivation is the fact that education is undergoing rapid digitalization
- Innovative instruction techniques, anywhere anytime any size learning, flipped classroom approach, blended learning environments, are the need of the hour

Language Learning through Association Capturing (1/2)

Need for association capturing

- Associative networks help not only in understanding new knowledge but also to retain it firmly in the mind
- Vocabulary plays an important role in all the competencies of language learning, such as speaking, reading, and writing
- Methods of Vocabulary learning have moved from the behavioristic theory [3] of language learning to the modern methods of vocabulary learning [4]
- As per communicative theory, understanding a word involves committing to memory its form, capturing its meaning and knowing how & where to use it [5]
- Thus, the meaning of a word is understood as collections of associated concepts

Language Learning through Association Capturing (2/2)

Associative network based Wordnet for language learning

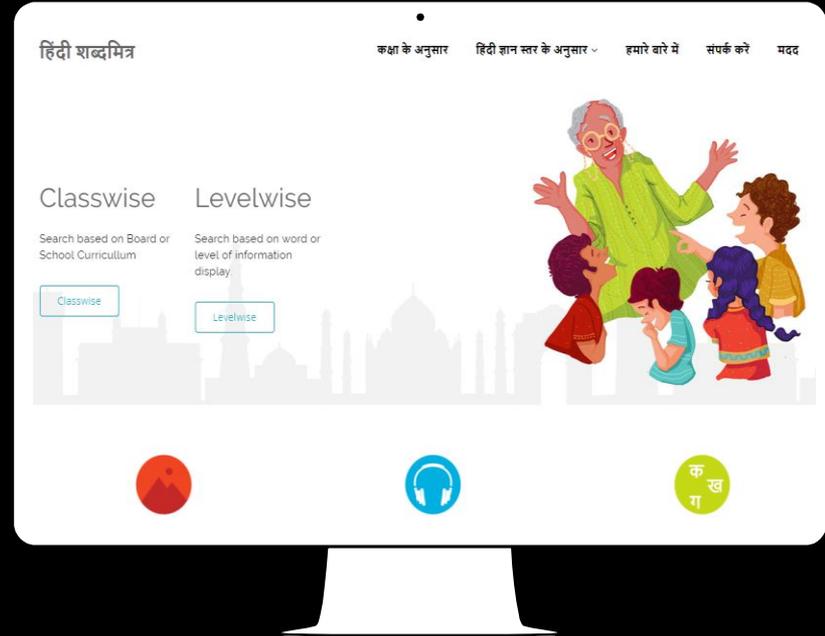
- The learning and teaching of vocabulary with associative networks is helpful as words are presented in semantically related group
- Meaning is captured and retained when they are taught in clusters. E.g., ‘a horse’ is a kind of ‘an animal’ (hypernymy relation); ‘uniform’ is opposite of ‘diverse’ (antonymy relation); ‘snore’ is a part of ‘sleep’ (entailment relation) ; ‘cultivate’ has as its object ‘land’ (argument relation)
- Such associations are not captured in conventional dictionaries but can only be provided by semantic networks like Wordnets
- Various tests have proven that amongst the plethora of web-based resources for language teaching and learning where WordNet has emerged as one of the most reliable, authentic and useful sites [6]

A Digital Aid

for Language Teaching
and Learning:

Hindi Shabdmitra

[http://www.cfilt.iitb.ac.in/
hindishabdmitra/](http://www.cfilt.iitb.ac.in/hindishabdmitra/)



Hindi Shabdmitra

- A comprehensive digital aid for Hindi language teaching and learning
- For promoting standard Hindi language learning
- Based on the psycholinguistic principles
- Vocabulary and grammar learning through images and audios
- Suitable for both formal and informal teaching environment
 - Formal: School set-up
 - Informal: Self learning
- Available in the form of web interface and mobile interface
- Hindi Wordnet has been transformed for pedagogical purpose in the form of an e-learning tool - ***Hindi Shabdmitra***

Multi-modal learning - a psychological aspect

- Human brain processes the inputs received from different senses to comprehend concepts
- Multi-modal learning leads to comprehensive learning of a concept with higher retention rate
- A multi-modal e-learning tool provides the learner a motivation for learning the second language along with the independence to learn at one's own pace
- In Digital aid, information is provided in the form of text as well as audio-visual inputs
- The textual information pertains to the gloss (original Hindi wordnet or simplified), word usage, synonyms, grammatical features, lexico-semantic relations, ontological information, *etc.*
- It also has audio pronunciations of words and pictures/illustration of concepts

Multi-layered presentation – an incremental learning approach

- The presented Digital aid is designed keeping in mind various aspects of language teaching and learning
- Selective information is rendered in every layer, depending upon the type/need of the user and his/her cognitive competence
- Level-wise information presentation – for informal setup (self pace learners)
- Class-wise information presentation – for formal setup (schools and institutes)



Class-Wise

हिंदी शब्दमित्र

कक्षा के अनुसार हिंदी ज्ञान स्तर के अनुसार हमारे बारे में संपर्क करें मदद

पात

CBSE
Class 8
Lesson 1

मृदुल
हरा
पात
गात
निद्रित
प्रसूष
मनोहर
फेरना
पुष्प
तंत्रालस

पात



NOUN

परिभाषा - पेड़-पौधों में होने वाला विशेषकर हरे रंग का वह पतला, हल्का अवयव जो उसकी टहनियों से निकलता है

वाक्य में प्रयोग - केले का पात बड़ा होता है।

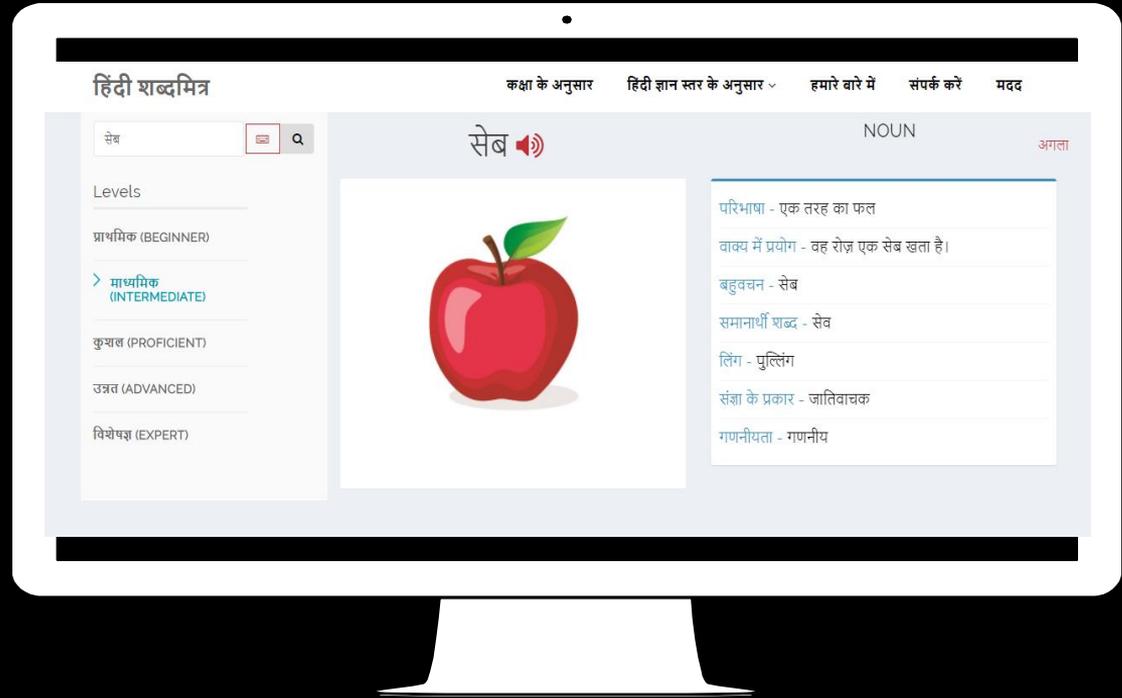
समानार्थी शब्द - पत्ता, पर्ण, पत्र, दल

लिंग - पुल्लिंग

एक तरह का - वनस्पति अंग

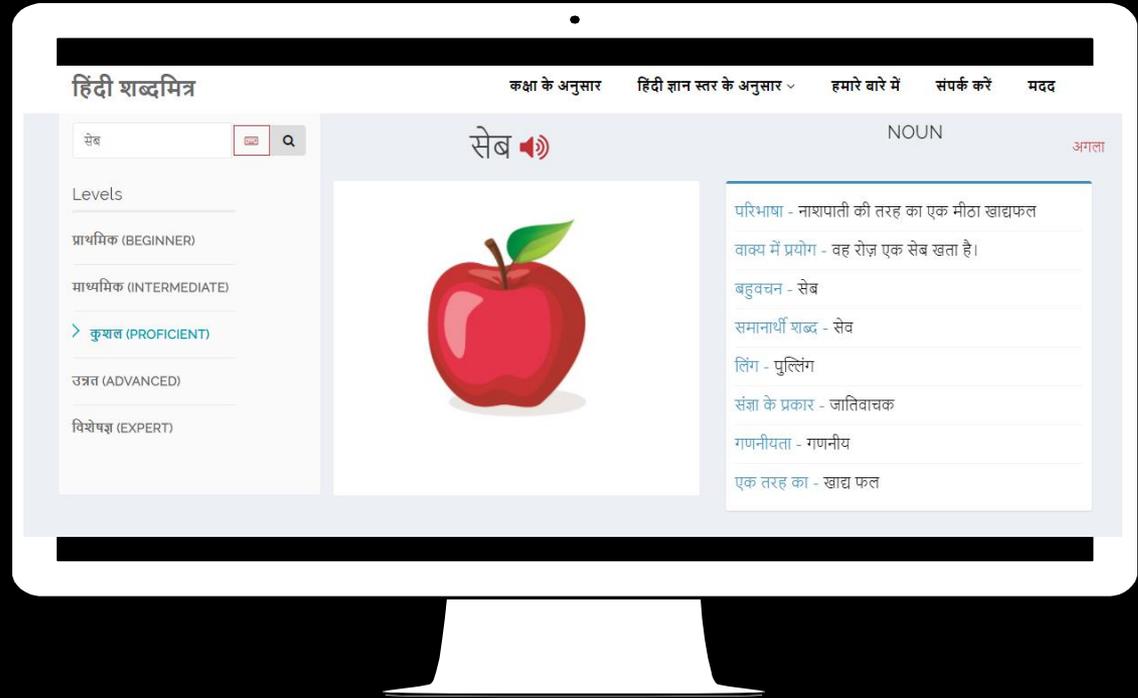
प्रकार - तैजपत्ता, पान, कौपल, तुलसीदल, पत्ती, कढ़ी पत्ता, बेलपत्र

Level-Wise Intermediate Level



Level-Wise

Proficient Level



The screenshot displays the Hindi Shabd Mitra website interface. At the top, the title "हिंदी शब्दमित्र" is visible, along with navigation links for "कक्षा के अनुसार", "हिंदी ज्ञान स्तर के अनुसार", "हमारे बारे में", "संपर्क करें", and "मदद". A search bar contains the word "सेब". Below the search bar, a sidebar lists proficiency levels: "Levels", "प्राथमिक (BEGINNER)", "माध्यमिक (INTERMEDIATE)", "कुशल (PROFICIENT)", "उन्नत (ADVANCED)", and "विशेषज्ञ (EXPERT)". The main content area features the word "सेब" with a speaker icon, a large illustration of a red apple, and the word type "NOUN". A list of definitions and examples is provided on the right, including "परिभाषा - नाशपाती की तरह का एक मीठा खाद्यफल", "वाक्य में प्रयोग - वह रोज़ एक सेब खता है।", "बहुवचन - सेब", "समानार्थी शब्द - सेव", "लिंग - पुल्लिंग", "संज्ञा के प्रकार - जातिवाचक", "गणनीयता - गणनीय", and "एक तरह का - खाद्य फल".

Process, Experiences and Lessons Learnt (1/5)

Word Collection

- The words are collected from Hindi textbooks by NCERT [7] which promotes standard Hindi language
- These books are followed by majority of schools across India and also in some schools in other countries
- The words which are not available in Hindi wordnet, but are present in textbooks, are collected and added in the database
- Examples: Proper Nouns [नागाजर्जु (*naagaarjuna*, Nagarjun)], Rhyming words [गमगम गमगम (*gamagama gamagama*, runs)], English words [लेमन (*leman*, lemon)], Idioms and Proverbs [मुँह में पानी भर आना (*muḥa meM paanii bhara aanaa*, mouth watering)], Game names [पकड़म-पकड़ाई (*pakaDama-pakaDaaai*, catch-catch), *etc.*

Process, Experiences and Lessons Learnt (2/5)

Gloss Simplification

- Hindi Wordnet was somewhat complex for the understanding of a language learner at a beginner stage
- Gloss and examples are simplified for level 1 and 2 for ease of understanding
- For a word हिम्मत (*himmat, courage*), the original Hindi wordnet gloss is: मन की वह दृढ़ता जो कोई बड़ा काम करने में प्रवृत्त करती है या जिसके कारण हम निडर होकर किसी खतरे आदि का सामना करते हैं (*mana kii vaha dRiDhataa jo kooi baDaaa kaama karane meM pravRitta karatii hai yaa jisake kaaraNa hama niDara hokara kisii khatare aadi kaa saamanaa karate haiM*, that perseverance of mind which motivates us to do some great work, or because of which we face fear and danger)
- A simplified gloss for the same concept is: मन की ताकत (*mana kii taakata, strength of mind*).

Process, Experiences and Lessons Learnt (3/5)

Picture Depiction

- In Hindi Wordnet, several concepts are hard to explain using the gloss
- A complex concept can be easily explained by a picture or an illustration
- Example: the concept of a word 'milk' in Hindi is explained as वह सफेद तरल पदार्थ जो स्तनपायी जीवों की मादा के स्तनों से निकलता है (*vaha sapheda tarala padaartha jo stanapaayii jiivoM kii maadaa ke stanom se nikalataa hai*, a white nutritious liquid secreted by mammals and used as food by human beings)
- Here, the illustration of a concept 'milk' helps in understanding the concept clearly



Process, Experiences and Lessons Learnt (4/5)

Grammatical Feature Marking

- The grammatical features of a given word are provided at each level of tool
- These features are marked by the lexicographers during the process of word collection and gloss simplification
- Each word is marked with the grammatical properties w.r.t. its POS category
- Examples:
 - Nouns are either countable or uncountable
 - Type of verbs are simple verb, conjunct verb, compound verb, causative verb
 - Kinds of Adverbs are manner, place, time and quantity
 - Kinds of adjectives

Process, Experiences and Lessons Learnt (5/5)

Audio Pronunciation

- Manually recording pronunciations for all the words is a tedious task
- These recording efforts are minimized by using text-to-speech (TTS) systems to automatically synthesize speech for most of the words

Field Test and User Feedback

- The prototype of the Digital Aid was demonstrated in three local schools
- The feedback was sought for content, ease of handling the application, classroom impact and overall experience by teachers and students
- Observed that the digital aid helped teachers in explaining concepts clearly with the help of images and simplified concepts even for primary language learner
- The aid assisted teachers in better classroom management, especially with the help of illustrations and reduced effort of reiterating the concepts for better retention
- The application has been improved based on the feedback received

Conclusion and Future Work

- A lexically rich resource, Hindi WordNet, is suitably modified and enhanced for developing a digital aid, Hindi Shabdmitra, for language teaching and learning
- It is a multi-modal multi-layered Hindi language learning aid which can be used for formal and informal learning environments
- The process followed, the experiences earned, the challenges faced and the lessons learnt are recorded in this paper
- Simplified concepts, grammatical features, illustrations and clear pronunciations has led to better classroom management and increased interest in leaning
- In future, Digital Aid can be expanded to the other Indian languages. Gamification and evaluation techniques will be incorporated

Acknowledgements

- The members of Center for Indian Language Technology (CFILT), IIT Bombay
- Research and development team of Hindi Shabdmitra, IIT Bombay
- The funding agency, Tata Center for Technology and Design (TCTD), IIT Bombay
- Bennett University, Greater Noida, India

References

- [1] Hindi Wordnet: <http://www.cfilt.iitb.ac.in/wordnet/webhwn/index.php>
- [2] Kenneth Ward Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational linguistics*, 16(1):22–29.
- [3] Mehmet Demirezen. 1988. Behaviorist theory and language learning. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 3(3).
- [4] Paul Nation and Jonathan Newton. 1997. 19 teaching vocabulary. *Second language vocabulary acquisition: A rationale for pedagogy*, page 238.
- [5] H Douglas Brown. 2000. Principles of language learning and teaching.
- [6] Amit C. Hiray. 2015. Teaching and Learning of EAP Vocabulary: A Web-based Integrative Approach at the Tertiary Level in India. *Ph.D. thesis*, HSS Dept., IIT Bombay.
- [7] NCERT: <http://ncert.nic.in> (National Council for Education Research and Training)

Thanks!

Any questions?

You can contact us at

- hanumantredkar@iitb.ac.in
- rajita.shukla38@gmail.com
- sandhya.singh@gmail.com
- jayasaraswati@gmail.com
- laxmi.kashyap@gmail.com
- Dipteshkanojia@gmail.com
- preethijb@gmail.com
- malhar@iitb.ac.in
- pb@cse.iitb.ac.in

