# Crowd-sourced Technical Texts can help Revitalise Indian Languages

**K. V. S. Prasad[1], Shafqat Mumtaz Virk[2], Miki Nishioka[3], C. A. G. Kaushik[4]**
[1]Department of Computer Science and Engineering, Chalmers University of Technology, Sweden.
[2]Språkbanken, Department of Swedish, University of Gothenburg, Sweden.
[3]Graduate School of Language and Culture, Osaka University, Japan.
[4]Independent Researcher, Gothenburg, Sweden.
prasad@chalmers.se, shafqat.virk@svenska.gu.se, dumas@lang.osaka-u.ac.jp, kaushikcag@gmail.com

**Abstract**

Many Indian language (IL) speakers use English words for all STEM (Science, Technology, Engineering and Mathematics) concepts, however elementary, ignoring the STEM vocabulary in IL textbooks up to high-school. People assume English is necessary, and ILs are unfit, for STEM and higher education generally. English and STEM competence also mark wealth, so parents now abandon first language (L1) schools for often woeful "English" ones even at primary level. So children learn everything poorly: L1, English and content. To reverse this collapse, people need to use L1 more broadly. This paper calls for IL STEM texts, crowd-sourced from STEM-trained IL-speakers, to seed such usage. We note how the texts would fit in the linguistic landscape. They would also be important new data for computational linguistics. STEM-trained people with rusty L1 writing, like us, will find that with the dictionaries and text online, they can write in L1—we comment on vocabulary and help from related languages. Crowd-sourced texts vary in quality, but they can help people to use L1 for STEM topics, and to realise that children learn content better in L1 than in bad English.
**Keywords:** crowd-sourcing, STEM, Indian languages, translations

## 1. Background and Overview

The Multilingual FrameNet Project (MLFN, 2017) is using translations of the popular TED talk (Robinson, 2006) for cross-linguistic comparisons of specific uses of words. We were to work (Virk and Prasad, 2018) with the Hindi, Kannada, and Telugu translations, but these had not yet been made. We translated the talk ourselves, finding that with help from the dictionaries and IL text online, even our rusty IL writing skills could convey factual non-literary content.

That Indians can access this talk only through English is typical. There is little writing in Indian languages (ILs) about STEM (science, technology, engineering, mathematics, and related subjects). This diminishes ILs, and harms IL speakers, as we shall see.

A body of online IL STEM[1] texts, added to regularly, will help revive ILs. The texts can be crowd-sourced from STEM-trained IL speakers, an effective solution that will also build IL communities.

## 2. Indian languages today

ILs are used in daily conversation, songs, films, television, newspapers, and to talk about politics, religion and philosophy. IL stories, poems and translations of literary works are published, if mostly in small circulation journals. School texts up to the final year are available in ILs (NCERT, 2012; TN Govt., 2011) for most subjects including science and mathematics, as are glossaries from CSTT(CSTT, 2018).

But ILs provide few college texts in STEM, and English is needed to study at university or get a job (Sharma, 2017; The Hindu, 2016), or just to read or publish. Even in other fields, English is often the primary language, as a visit to any bookshop will confirm.

Governments impose official ILs on employees and students, with examinations to be passed for promotion. Why do ILs need to be imposed? Because they have little academic prestige and no employment value in many desirable fields: they use English terms to deal with the simplest technical matter. All over India, any conversation with a professional (doctor, engineer, lawyer) is peppered with English words. The school books and glossaries cited earlier provide the needed STEM words in ILs, but the public rarely uses them, and the resulting conviction that STEM can only be spoken about in English devalues ILs, and harms all IL based education and even thinking.

## 3. The costs of a devalued L1

We look at the cases of first language (L1) or English as the medium of instruction in school, and then at an L1 being actively dismantled by its own speakers.

### 3.1. First language (L1) medium

Some students in rural areas, e.g. in Karnataka, have good L1 and STEM skills, but avoid higher STEM courses for fear of the English medium these entail (Sharma, 2017). Poor English preparation has robbed these students of their preferred field of study.

Many students from Tamil medium schools do well at STEM and enter engineering colleges, but struggle there because their school English is poor(The Hindu, 2016). Even if their college lecturers speak to them in Tamil, the English terminology is daunting. There are few Tamil books or other sources at college level, so they can fail even at their favourite STEM subjects. If they pass, English still threatens their job prospects.

These cases are bad, but redeemable, since the students are actually well educated, with good L1 and STEM skills. All they need is (1) better English in high school, and (2) L1 STEM texts to help in college:

[1]We use STEM both as a noun, meaning the subjects, and as an adjective, meaning "pertaining to those subjects".

1. Remedial English and STEM English. This is a matter of organisation, the responsibility of the sending school, receiving college, and state government. There are open access resources to help.

2. Make L1 translations of key papers and parts of books, as well as original material and L1-English bilingual glossaries, freely available online. Thousands of STEM students and professionals speak ILs, so this material can be crowd-sourced. Accumulated over time, it will help successive batches of students. Some source material will be copyrighted, but much is under free licenses.

### 3.2. English medium

The outcome depends on the quality of English.

1. **Poor English: serious damage.** Having seen problems with L1 as medium, and the need for English to get jobs, many parents pay money they can ill afford to send their children to "English medium" schools. Some states might move much instruction to English (TOI Edit., 2017).

   But the results may be disastrous (Mody, 2017), because English is spoken at best very badly by many teachers. Bilingual instruction is not possible: "teachers lack proficiency, never mind fluency, in English. So, their classrooms are not bilingual, just badly mixed up" (Mody, 2017). The result is that children learn nothing: not English, not L1, and not the subjects being taught. Hopefully, the children retain at least spoken L1, from which recovery can start.

2. **Good English: success and alienation.** Good English medium instruction from childhood will remain something for the few. These often do well with education and jobs, but other problems lurk. For all but a very few, Indian English (IE) is restricted to prose, with poetry and song being the preserve of ILs. So IE is not a full language. Those who speak IE do best if they have a full L1, with its literature and songs, perhaps even able to talk about their work in L1 with minimum code-switching. Otherwise, they might end up culturally impoverished in both L1 and English.

### 3.3. English devalues L1 by invasion

The rush to English can take absurd forms even outside education, with L1 speakers making their own language more "English", and actively damaging it.

Telugu speech now replaces many Telugu words by English ones. So "blood" in Telugu is no longer `raktaṃ` but `blaḍ`. Other examples among many are: Mummy, Daddy, book, books (the plural is separately imported), water, rice, oil, door, cotton, life, food, dog, cat, moon, soul, body, week, month, daily, against, common, open, enquire, and all numbers. Mostly nouns, but no part of speech is immune.

These imports are not like "bus","radio" or "telephone" which arrived along with the object. Nor are they advanced terms with no Telugu equivalent; these are unnecessary replacements of basic words. They do not help in learning English (the imports adapt to Telugu phonology and grammar), but impoverish vocabulary and make Telugu word games and songs harder.

Telugu once took `rakta` "of blood" and `pōṭu` "thrust", and made the transparent `raktapōṭu` "blood pressure". Now it has "b.p.", an opaque name. Since Telugu "blood" is now `blaḍ`, the loss of `raktapōṭu` is itself no longer a loss of transparency (perhaps this kind of loss should be named the *raktapotu syndrome*).

## 4. Current efforts have failed ILs

We believe a high-functioning L1 is essential to a person's cognitive, social and psychological well-being. Everyone needs English for international contacts in STEM and for business, but most Europeans, for instance, use second-language English for this, while at home doing everything in L1, including talking to doctors, mechanics, lawyers and bankers.

But seventy years after independence, India is more dependent on English than ever, and does not have this privilege. No Indian language, most of which are larger than most European languages, can manage STEM without English words, and large numbers of people are shut out of participating in STEM activities (Sharma, 2017). We think India's language policies have failed, utterly—they have produced the current situation. The example below should give pause.

**Over-centralisation.** The 2017 ASER report (ASER, 2017) notes in passing how far language centralisation has gone. They found that oral rehydration solution (O.R.S.) packets are available in only Hindi and English—across all states in India. (To assess whether people could read and understand written instructions, the O.R.S. text was created in all 13 languages of the ASER survey). We can but repeat: to enable development, and protect life, respect L1.

**Books are not enough.** Textbooks, and some popular books and magazines on STEM topics are available in ILs. Government has tried to take STEM to the people in their own languages[2], and individuals have published popular science books in, for example, Telugu and Kannada (Verne, 2017; Vemuri, 2017; Hegde, 2017). Why are these efforts not enough?

Because without follow-up, books and magazines leave you high and dry. Also, no fixed set of texts can meet all needs. If you're going to Italy, a book on China does not help. English readers simply look up Italy in Wikipedia or other regularly updated websites. We would like to give IL readers the same luxury.

## 5. Clarification of our goals

The long-term goal is a high-functioning L1 in everyday speech, overturning the conviction that the sim-

[2]E.g., Two popular science periodicals in Hindi(Pragati, 2017; Homi Bhabha Centre for Science Education, 2017) publish articles accessible after registering as a user and logging in. The regional languages are looked after by state governments, but do not have the same level of resources.

plest STEM matter is the exclusive preserve of English. For this, IL-STEM vocabulary has to be used, regularly, in L1 speech and writing. Speakers should be able to replace English words by L1 words unselfconsciously, both in conversations about technical matter, and in college STEM texts (these latter might occasionally have to invent new L1 STEM vocabulary).

The reward will be the extension of L1 immediacy of understanding and expression to STEM matters. Children can see STEM being discussed in L1, both in real life and on social media, without reaching quite so often for an English word. The L1 words from their science and mathematics books will become real, and the word building and flow of L1 will work here too. With good L1, and STEM acquired through L1, the addition of English will become a less critical second language matter.

These goals are but dreams just now. As seeds, and continuing support for such development of ILs, we suggest crowd-sourced open access articles.

## 6. Seed: crowd-sourced STEM texts

Since crowd-sourcing is *crowd* based, it can produce the articles and websites mentioned above, in various languages, independent of government action or policy other than access to the internet. Just as cable television opened many different language channels, the internet can be used to revive ILs. People can now improve their own languages. How well any language group does will depend on how well their STEM trained speakers respond to this need.

We now outline a plan to build a body of STEM writing in L1, and say why it is feasible.

**Hosting.** The host for the contributed texts should provide minimal supervision and editing, but require proper acknowledgment of sources. It could focus on a particular L1 and provide aids for writers: links to online dictionaries, wiktionaries and other resources (several noted below), and previous contributions. It could provide tools for L1 such as parsers, treebanks, and wordnets, and forums where subject experts can collaborate with L1 experts, or where students can request L1 texts on specific STEM topics.

**Amateur writers as contributors.** Contributors need only be fluent L1 speakers who can write about the subject at hand (in English), and are willing to try to do so in L1. It does not matter if they last wrote in L1 long ago, if at all, or if their first L1 writing is difficult and slow, and less than perfect. The goal is only to convey the content of their ideas or of the original text they are translating.

**Low barriers to entry.** Using only known contributors might offer better quality, but the first requirement here is to get going. Style will come with practice, and texts can be improved.

The point is to encourage as many writers as possible, of both original texts and translations. The latter help build parallel corpora and provide texts outside the scope of the writer's immediate expertise.

**Won't machine translation (MT) help?** Yes, but first we need STEM texts in ILs. Statistical MT (SMT), such as used in Google's translation system, needs parallel corpora; with interest in IL technical texts decreasing, as it is now, there will be fewer texts for SMT to feed on, and it will be a self-fulfilling prophecy that MT fills social needs. Our goal of strengthening ILs breaks this downward spiral.

SMT is impressive, but less so for Telugu than for Hindi, for instance, presumably because there is less parallel corpus data for Telugu than for Hindi. So the crowd-sourcing we call for will be useful even if only to improve SMT systems[3].

It does appear that even the best MT systems still need a human to clean up the output, strengthening our argument that MT is only a support for languages people actually care about.

## 7. Notes on our novice L1 writings

We are IL speakers who write about STEM in English. Our L1 writing skills are rusty, and we have not read much L1 literature. In our L1 writing, we hope only to convey content. So we ourselves fit the minimum profile of the contributors we seek.

We found that with the dictionaries and text now online, amateur IL writing is quite feasible. This is anecdotal, but will hopefully encourage enough IL STEM contributions to generate more systematic data.

**Notation.** We use standard notation for the IL sounds. The palatalised spirant is written š. Retroflexion is shown by a dot under the letter; ṛ, a flap, is limited to Hindi and Urdu; ṇ and ṣ to Sanskritised words; and ḷ to Telugu and Kannada. Aspirated stops are shown thus: kʰ. A macron over a vowel denotes a long vowel, and ~, nasalisation. In Hindi and Urdu, e and o are always long, so the macron is dropped. ñ is the nasal homorganic with the following consonant.

### 7.1. Resources

Listed for Hindi, but apply similarly to other ILs.

**Dictionaries.** Wiktionary(Wiktionary, 2017b) uses English as the meta-language[4] and offers etymologies: (Wiktionary, 2017a) has sub-pages for words borrowed from specific languages. Over a thousand Perso-Arabic content words are listed for Hindi; Urdu would have more. This is a partial list: e.g., होशियार hošiyār "clever" is listed, but not होश hoš "consciousness".

Collins English-Hindi dictionary (Collins, 2017) is good, but limited in scope; the English example sentences are excellent, but there are no Hindi ones. Classic dictionaries are available from (Univ. of Chicago, 2017). Shabdkosh (Shabdkosh, 2017) is useful for throwing up many possible synonyms, but offers no contexts and examples to choose between them.

[3]The alternative to SMT, rule based MT, goes back for ILs at least to the Akshar Bharati group (Akshar Bharati et al., 1996) since the early 1990's. The Anusaaraka translator (Anusaaraka, 2017) is a testament both to the quality of their work, and to the difficulty of rule based MT.

[4](Hindi-Wiktionary, 2017) has Hindi descriptions.

**MT.** Google's translate is becoming very good indeed, but still needs cleanup, and does not distinguish between forms of Hindi[5].

Urdu and Hindi "share the same grammar and most of the basic vocabulary of everyday speech" (Flagship, 2012; Prasad and Virk, 2012); and (Bhat et al., 2016) says they are different standard registers or literary styles of the same language. So translation between the two should be shallow transfer, involving almost only lexical substitution. Apertium (Apertium, 2017) is a tool that does such jobs. Our Hindi translation of the Robinson talk begin by feeding an existing Urdu translation (Hassan and Anjum, 2006) into Apertium, and manually cleaning up the output.

### 7.2. Cultural connotations

The problems of translating across "unbridgeable cultural differences" are dealt with by (Prattipati, 2017) in the context of Bible translation to Telugu. Perhaps surprisingly, such problems can appear in technical matters too.

Mathematics has cultural foundations as (Raju, 2007) points out. "Proof" is translated to Hindi as प्रमाण `pramāṇ`, but in mathematics the former nowadays means logical deduction, whereas "Indian philosophy considered empirical proof `pratyakṣa` as more reliable than logical inference" (Raju, 2013). So the translation may have very different connotations, particularly for the reader in touch with their linguistic roots.

**Sanskrit for STEM?** Following on from that last remark, we note that there seem to be few STEM-trained people who can read the Sanskrit scientific and mathematical literature, and it seems to us that one long term goal is to build up such a community. Whether or not one agrees with Pollock in most of his paper (Pollock, 2011), it does seem that he is right that colonialism did "[...] render the literary past unreadable to most Indians".

### 7.3. Telugu and Kannada

**Agglutination.** Telugu and Kannada are agglutinative languages, so it is easy to produce a word that is in no dictionary. The question for the unsure writer is how to check that the word is acceptable. The first step is to search for the new word and see if it appears in other texts on the web. If it does occur, it is important to judge if the source is trustworthy.

If the word is not found on the web, the next step is to search for variations. So if వచ్చినట్టున్నాడా `vaccinaṭṭunnāḍā` "does it look like he has come?" is not found, then search for `vaccinaṭṭunnāḍu` "it looks like he has come" or `koṭṭinaṭṭunnāḍu` "it looks like he has hit", etc., to confirm that a nearby structure is in use. Native speaker judgement can be trusted to say that if "it looks like he has come" appears, then the proposed "does it look like he has come?" is acceptable.

**Vowel Harmony.** This is a feature of Telugu but not of Kannada. It means that in Telugu it is sometimes not clear what vowel one is using in speech. So కఠిచి `karici` "having bitten" can also be written and spoken `karaci`. Again, the solution is to search for the various possibilities and see what turns up, which is more popular, and so on.

### 7.4. Related languages

Telugu and Kannada are closely related languages. A translation of a sentence from one to the other often preserves morpheme order. This suggests that these languages make a good candidate pair for Apertium.

A novice IL writer will find themselves learning more about their L1, and so this paper ventures to comment on this experience. A traditional way to strengthen L1 is to learn languages close to L1; e.g., it used to be that many Kannada speakers also learned either Telugu or Tamil—useful for, say, music lyrics. But now, the three language formula means neighbours can often communicate only via Hindi or English. Learning these, or indeed any other, languages is good, but losing languages close to L1 is not.

### 7.5. Sanskrit borrowings

Since Sanskrit is a very important vocabularly resource for most ILs, it is important to understand how it fits in with any given L1.

For example, Telugu and Kannada are full of Sanskrit words commonly used, with no air of formality. As in many other ILs, Sanskrit has long been digested. The same word can even appear in original and many assimilated forms. Telugu has `nidra` "sleep" as in Sanskrit, but also `niddara` and `nidura`. Kannada has both `mūrti` "form" and `mūruti`, and so on.

Both the original form and the assimilated ones sit comfortably in speech, and in word building. "Go to sleep" can be `nidrapō` or `nidurapō` in Telugu[6]. But note that in Hindi, निद्रा `nidrā` can make निद्रायमान `nidrāymān` "one who is asleep", but नींद `nīnd`, the assimilated form of the Sanskrit word, cannot.

**Word-building.** This is important in STEM texts. E.g., "add" can be translated into Hindi as जोड़ना `jōṛnā` or योग करना `yōg karnā`, but only the latter Sanskrit form can make योगात्मक `yōgātmak`, "additive".

## 8. Summary

We have seen that currently L1s in India are devalued, and English, however poor, is seen as the route to progress. The context of the paper is a dream of restoring ILs to the status of full languages, used daily also for STEM and other technical subjects, and thus paving the way to fuller individual development.

1. Our primary contribution is a call to crowd-source STEM articles in ILs to complement existing IL

[5] "Hindi" covers quite different dialects, including Hindustani and "shuddh" Hindi. It thus has multiple *forms* (Kachru, 2006), and Standard Hindi is hard to define.

[6] Note that both `nidrapō` and `raktapōṭu` combine a native Dravidian word with a Sanskrit word, illustrating the degree of comfort with Sanskrit.

textbooks and glossaries, and keep up with new theories and technologies. These articles can help revitalise ILs and allow their speakers to comfortably navigate new worlds of ideas instead of being forced to do so in a foreign language.

If the new texts are translations, they will also improve the performance of MT for their language.

2. As STEM-trained IL speakers with rusty IL writing skills, we found in our writing efforts that the dictionaries and texts already online helped us write in L1 fairly confidently. So we believe our call is workable. In Sec. 7, we noted several factors of which the novice IL writer is likely to experience parallels.

3. We suggest that the star-shaped landscape of ILs today (interstate communication only via Hindi or English) should be softened to allow people to learn the languages close to their L1, a traditional way to strengthen L1, and allow language networks develop more naturally. Life critical information should be available in L1.

   A related decentralising idea is to develop tools to develop translators, such as Apertium, and other tools to work with related languages (such as Telugu and Kannada).

4. To go deeper into ILs could mean rediscovering Indian understandings of concepts such as proof. It would be useful to have more STEM-trained people who can read the Sanskrit STEM literature.

**Conclusion.** Further possibilities will reveal themselves as we go along. Knowledge grows only by sharing. Thanks to poor language development, India has denied its people full development, and grossly underutilized its human intellectual capital. We think crowd-sourcing can help set this right. It is a quick, cheap and inclusive way of tapping into existing potential, improving both understanding and L1 skills, and building a community of L1 writers.

**Acknowledgement.** We thank the referees for their comments, which focussed the paper and helped us remove extraneous material.

## 9. Bibliographical References

Akshar Bharati, Chaitanya, V., and Sangal, R. (1996). *Natural language processing: A Paninian perspective*. New Delhi: Prentice Hall of India.

Anusaaraka. (2017). Anusaaraka home page. https://anusaaraka.iiit.ac.in.

Apertium. (2017). Apertium wiki main page. http://wiki.apertium.org/wiki/Main_Page.

ASER. (2017). Annual Status of Education Report. www.asercentre.org.

Bhat, R. A., Bhat, I. A., Jain, N., and Sharma, D. M. (2016). A House United: Bridging the Script and Lexical Barrier between Hindi and Urdu. In *COLING*, pages 397–408. ACL.

Collins. (2017). English-Hindi dictionary. https://www.collinsdictionary.com/dictionary/english-hindi.

CSTT. (2018). Commission for Scientific and Technical Terminology, Main Page.

Flagship. (2012). *Undergraduate program and resource center for Hindi-Urdu at the University of Texas at Austin.* http://hindiurduflagship.org/about/two-languages-or-one/.

Hassan, S. and Anjum, U. (2006). Urdu translation of: Ken Robinson, Do schools kill creativity? TED profiles 133520 and 934090.

Hegde, N. (2017). Speaking science from the Kannada stage. http://www.thehindu.com/books/speaking-science-from-the-kannada-stage/article19974531.ece.

Hindi-Wiktionary. (2017). hindi vikshanari. https://hi.wiktionary.org/wiki/ .

Homi Bhabha Centre for Science Education. (2017). homi bhabha vijnan shiksha kendra. http://ehindi.hbcse.tifr.res.in.

Kachru, Y. (2006). *Hindi*. Philadelphia: John Benjamins Publ. Co. London Oriental and African Language Library.

MLFN. (2017). Multilingual FrameNet Project. framenet.icsi.berkeley.edu.

Mody, A. (2017). Lost for words: Why Tamil Nadu's shift to English medium instruction is not helping children. *scroll.in*. 17 Aug 2017.

NCERT. (2012). *School textbooks*. New Delhi: National Council for Educational Research and Training.

Pollock, S. (2011). Crisis in the Classics. *Social Research: An International Quarterly*, 78(1):21--48. India's World, a special issue. www.socres.org.

Pragati, V. (2017). vijnan pragati. http://www.niscair.res.in/sciencecommunication/Popularization%20of%20Science/vigyan0.asp.

Prasad, K. V. S. and Virk, S. (2012). Computational evidence that Hindi and Urdu share a grammar but not the lexicon. In *3rd Workshop on South and Southeast Asian Natural Language Processing (SANLP), collocated with COLING 12.*

Prattipati, M. (2017). The Holy Register: Its Equivalence and Strategies. *Translation Today*, 11(1). http://www.ntm.org.in/download/ttvol/volume11-1/Art_4.pdf.

Raju, C. K. (2007). *Cultural Foundations of Mathematics*. Pearson Longman.

Raju, C. K. (2013). Teaching mathematics with a different philosophy, part 1: Formal mathematics as biased metaphysics. https://arxiv.org/ftp/arxiv/papers/1312/1312.2099.pdf.

Robinson, K. (2006). Do schools kill creativity? *TED: Ideas worth spreading*.

www.ted.com/talks/ken_robinson_says_schools_kill_creativity/up-next.

Shabdkosh. (2017). English to Hindi/Telugu/Kannada Dictionary. shabdkosh.com.

Sharma, A. S. K. V. S. (2017). *Challenges of Communicating Science in Regional Languages: Experiments in Kannada.* Springer. *Bridging the Communication Gap in Science and Technology*, Pallava Bagla and V. V. Binoy (eds.).

The Hindu. (2016). Language barrier puts students from Tamil medium in a fix. *Chennai: The Hindu, 13 July 2016*, July. Vasudha Venugopal and B. Aravind Kumar.

TN Govt. (2011). *School textbooks (Telugu and Kannada)*. Government of Tamilnadu, Department of School Education.

TOI Edit. (2017). Reform schools: Andhra Pradesh government's bold step in making English medium compulsory heralds change. *Times of India, 20 January 2017.*

Univ. of Chicago. (2017). Digital South Asia Library. dsalsrv02.uchicago.edu/dictionaries/.

Vemuri, V. (2017). Emi endukani. http://web.cs.ucdavis.edu/~vemuri/TeluguBooks.htm.

Verne, J. (2017). patalaniki prayanam. Telugu translation by Vaddadi Srinivas Chakravarthy of "Journey to the Centre of the Earth".

Virk, S. M. and Prasad, K. V. S. (2018). Towards Hindi/Urdu FrameNets via the Multilingual FrameNet. In *Proceedings of the International FrameNet Workshop (IFNW)*. To appear.

Wiktionary. (2017a). Category: Hindi terms by etymology. https://en.wiktionary.org/wiki/Category:Hindi_terms_by_etymology.

Wiktionary. (2017b). Wiktionary, the free dictionary. https://en.wiktionary.org/wiki/Wiktionary:Main_Page.