

# Democratizing Translation Using Digital Methods

**Gurumurthy, K | guru@itforchange.net**

**Anand, D | anand@itforchange.net**

**Karthik, K | karthik@itforchange.net**

K. Gurumurthy is the Founder and Director of IT for Change, Bangalore. His areas of expertise include ICT integration in school education, pre-service teacher education, in-service teacher education and policy advocacy. He also works in the areas of school leadership and free and open digital technologies.

D. Anand is a consultant with IT for Change. His work focuses on meaningful appropriation of digital technologies for social change and reflects his interest in multilingual approach to language learning, as well as resource and activity-rich learning environments in the classroom and beyond.

K. Karthik is a Programme Associate in IT for Change. He is interested in the intersection of digital technology with gender, adolescence issues and education. He is also interested in Kannada poetry-writing

**Key Words:** Digital translation, MediaWiki, Possible meaning, Digital terminology data banks, Translation memories, Human Aided Machine Translation (HAMT), WhatsApp, Online dictionaries, Web repositories, Video conferencing.

## Abstract

Translation is seen as a highly-skilled professional activity, which limits the number of practitioners and hence its reach. Three members of the NGO IT for Change that works with government schools attempted to translate a text from English to Kannada, using easily accessible digital tools. The NGO also supported teams of student-teachers and professional teachers in using these tools for translating education materials. The aim was to make the translation a participatory exercise "by teachers, of teachers, for teachers", and improve the availability of education resources in Kannada.

## Resource Gap

The volume of materials (articles, papers, media) in English far exceeds that of Indian languages. Wikipedia has over 6 million articles in English while Kannada has less than 30,000 (<0.5 per cent). Hindi, at 140000+, has the highest number of articles amongst Indian languages. This can make not knowing English a disadvantage, even though factors such as one's ability (text fluency) and desire to read also play a part in mitigating this disadvantage. English, however, is spoken by around 10 per cent of Indians (List of Wikipedias, n.d.). If we could translate the English articles into a local language content, they could reach many more people. The availability of contextual reading materials is also vital in developing an interest in reading in young minds. Translation can provide greater access to written literature as well as oral resources; and translation of articles on education-related topics would help teachers and teacher educators to broaden and deepen their understanding of education. It could also promote learning in the target language as well as teacher development.

This article discusses how a three-member team from IT for Change (ITfC) used digital MediaWiki platform technologies (MediaWiki, n.d.) for a simpler and quicker translation. Wikipedia content was used for translation because it is an open-source, under the "Creative Commons" licensing, which allows for its content to be modified and derivatives to be produced.

## Complexity of Trans-Creation

Translation cannot merely replace words from one language with the equivalent

words from the target language (meta phrase). In translation, one must look for words with potential for bearing possible meanings (Sarukkai, 2013), or affective instead of formal equivalence (Shivaprakash, 2020). A translated text must bring the source text closer to the local culture. Engaging in the process of translation helps readers understand and learn from the culture of the people who speak the source language (SL). It also introduces new ideas into the target language (TL). Translated resources must be crisp, not elaborate explanations or commentaries on the source text.

Translation is a complex process; it requires several skills. The translator must be competent in reading and writing in the SL as well as the TL, and first understand the content in the SL and then write it in the context of the TL.

Translation requires excellent vocabulary, knowledge of phrases and idioms in both languages to reflect cultural nuances in translation, as well as a reasonably good knowledge of the domain. The challenge is to capture the essence of the article so that a TL reader can have the same experience as a SL reader, yet situated in their own context. Given these complexities, 'translation' is sometimes termed as 'trans-creation'.

## Conventional Translation (Pre-Digital)

The conventional translation is a manual process. The translator tries to get equivalent words and phrases for SL in the TL, referring to bilingual dictionaries and thesauruses if required. The process of translation and error correction can be time-consuming and effort intensive. This method severely limits the number of people who can translate.

## Translation Using Digital Methods: APU Publications

Digital technologies used in translation include Human Aided Machine Translation (HAMT) such as Google Translate and Machine Aided Human Translation (MAHT) such as online dictionaries, which takes lesser time to consult compared to printed dictionaries. Digital terminology data banks that list special-purpose terminology enhance translation efficiency and accuracy, as well as translation memories, as previously translated texts are available digitally (Bhattacharyya, 2004). In MAHT, the translation is done by humans, unlike HAMT.

A three-member team from ITfC worked on the translation using simple, easily accessible desktop and cloud-based technologies to translate English articles (SL) from APU journals into Kannada (TL). The members of the team, work in school education projects and are not professional translators.

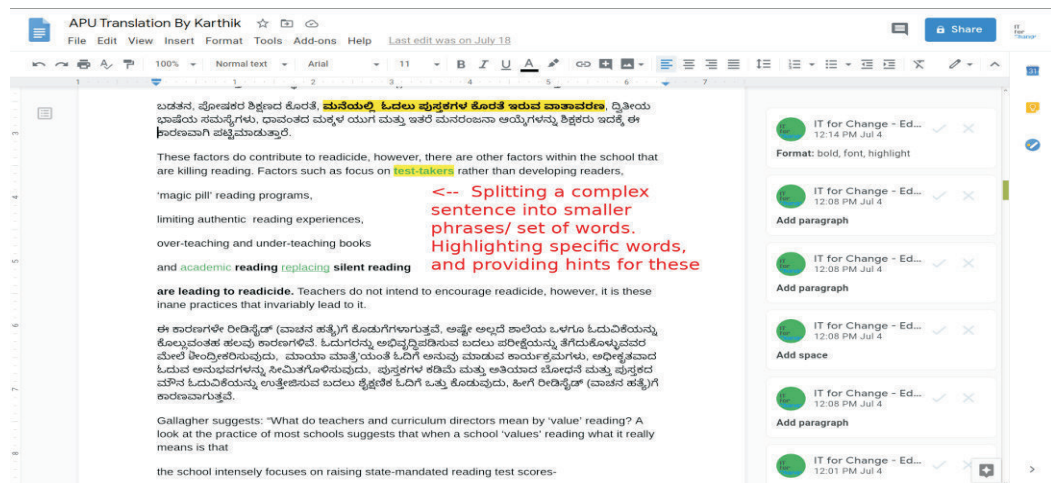
## Translation Process

The translator (T1 or T2) completed the initial translation of an article with the help of an online machine-translation tool such as Google Translate and *ShabdKosh*, and web repositories such as *Kanaja* and *Baraha*. They also consulted mono and bilingual digital dictionaries available online as well as offline on a computer or a mobile phone and listed the different meanings of a word/phrase identified through a simple "search/find" process. Oft-repeated terms used in education were stored in a simple tabular format on another cloud document on Google Sheets, to allow for easy reference. The team simultaneously accessed both dictionaries as well as Google Sheets.

The translator made paragraph-wise sections of the document and provided the TL translation just below the SL paragraph. The digital document therefore became a bilingual document, with the translated section available right below the original, for easy comparison (Screenshot1). This document, with full editing rights given to all members, was uploaded as a cloud document on Google Docs. A common Gmail-id was used by the team members to access it.

### Screenshot 1. A Paragraph Translated into TL

(Below the Relevant SL Section; Read the Comment in the Image. The "track change" is indicated by the font colour in Green, in Contrast to the Regular Text in Black.)



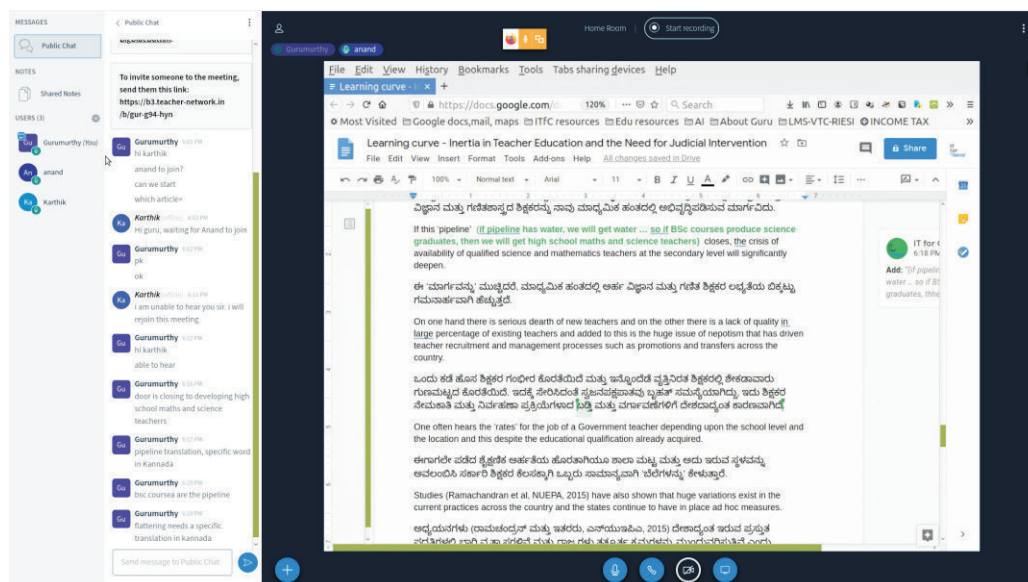
T3 "broke" complex sentences in the source text into meaning-holding units, adding punctuation marks such as "," or "-" or line-breaks for easier comprehension. T3 used the "track changes" feature of the document editor, so that his additions, modifications and deletions were highlighted. This allowed T1 and T2 to identify the original text and the changes made by T3.

Once the translation was complete, the team met on the free and open-source BigBlueButton video conferencing/webinar platform (like Google Meet or Zoom) for a collaborative review of the translation (Screenshot 2). The bilingual document was shared on the platform using the screen-share feature

to help the three members work on the same section simultaneously. All three translators also simultaneously edited the Google document on their computers.

T1 and T2, who worked on the TL text, did direct edits for real-time comprehension of the TL document and changes during the discussions. T1 read out the content in the TL, and T2 and T3 simultaneously assessed the quality of the translation. T2 reviewed the flow of the TL text and focused on SL-TL meaning equivalence. T3 checked if the spirit of the SL content had been adequately captured in the TL, and on its appropriateness in the education context. T2 made corrections, as required.

**Screenshot 2. Concurrent Editing on the Video-Conferencing Platform**  
(BigBlueButton – to Review and Revise the Translation)

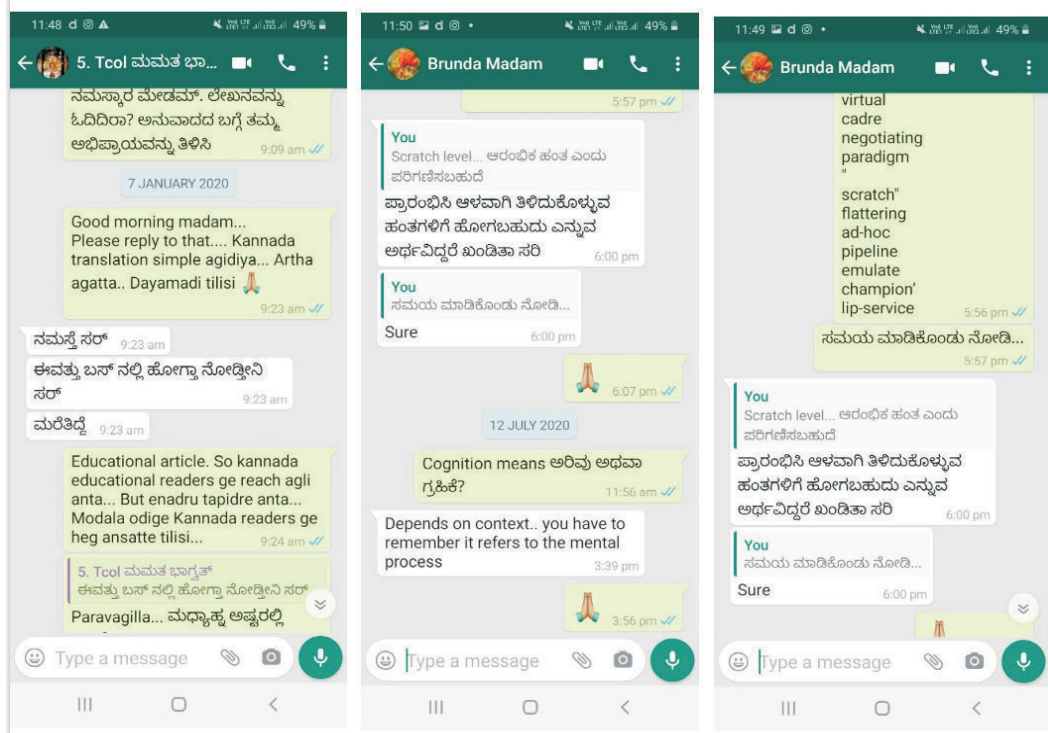


The translation process usually generates multiple interpretations of the SL as well as multiple expressions in the TL. The decision to choose a particular word or a phrase in TL depended on "maximum meaning", according to Sarukkai (2013). Whenever doubts or debates could not be resolved, the Kannada and English

Wikipedia pages were accessed to find equivalent words for the terms being debated (English Wikipedia, n.d). These dialogues helped to arrive at a consensus on the interpretation of the phrase in SL and/or the word/phrase to be used in the TL. Then the words/sentences were edited and revised by T2. Where the team

did not agree, T1 sought help from external experts by sharing the word or phrase or sentence in both SL and TL through asynchronous social media (WhatsApp) (Screenshot 3).

**Screenshot 3. WhatsApp helpline**



Once the draft translation was complete, T1 shared it with a friend in the Kannada teachers group over mail, to seek their review and comments. Since the translation was aimed at providing resources for the Kannada language/Kannada medium teachers, their review could help assess if that purpose had been achieved. After that, the translated article was mailed to APU for publishing on their website.

The subsequent step could have been to host the article on a Wiki, instead of (or along with) a static web page, so that readers/teachers could further refine/improve the translation. No translation can ever claim to be complete; this is true for any material creation process. We were able to implement this Wiki process of "infinitely iterative

improvement" in another translation project of ITfC. Our "Wikipedia content translation" (Content Translation, n.d) project was a practicum for student teachers of two teacher education institutions where we taught the "ICT Integrated Learning" course.

## Translating to Kannada Wikipedia

Papert and Harel (1991) refer to constructivism as "learning to create and creating to learn". Similarly, NCERT (2013) discusses "connecting and learning" and "creating and learning" as essential themes for integrating ICT into teaching-learning; these are equally relevant to translation. By connecting people, one can facilitate collaboration amongst

diverse participants. Digital tools enable the creation of materials in the TL, which can by itself be a developmental process. Wikipedia is an example of the confluence of these "connecting" and "creating" themes. Online encyclopaedias allow anyone with a digital device and internet connectivity to add to its repository. Several editors can edit the same page, across space and time.

We asked students to make two-member teams, with one member who was comfortable in English and the other, in Kannada. The former had to try to understand and interpret the article on the English Wikipedia page and that the latter had to write the translated article on the Kannada Wikipedia page (Kannada Wikipedia, n.d). Alone, each would have found this task impossible; together, they could attempt it. We taught students to use HAMT; they created the initial version of the English Wikipedia page in the Kannada Wikipedia, using the auto-translate option available in the Wikipedia content translator. We also taught them to use the voice to text feature, available on their digital devices such as google doc, where the team could read the SL, orally translate into text in the TL, which the software converted to text. This text was copy-pasted into the Kannada Wiki page.

The MediaWiki approach allows continuous refinement, instead of having to get it perfect the first time. This makes it easier to begin translation and then improve it gradually. Further, with this approach, the possibilities of involving teachers in translation on a large scale becomes possible. ITfC's Karnataka Open Educational Resources (KOER Project), which has both English and Kannada MediaWiki repositories, adopted this approach. Teachers and ITfC members created content on either repository or translated the content into the other repository, at different points in time.

HAMT tools often capture our translations and refine them, making machine translations more reliable over time. For instance, Google Translate uses the translations of its users to improve its translation engine. Regular use of online translation tools can refine both the tools and their associated repositories. These tools are examples of artificial intelligence (rather machine learning) that are unarguably beneficial to humanity. The ownership of the translation, however, should be public rather than private, for it to fully benefit society and not held ransom to the owners' commercial interests

As HAMT's accuracy increases, translation can become automated. However, HAMT will be more error-prone for complex articles, like many in the education domain. MAHT, however, will continue to be relevant. Thus, while the use of digital tools and methods can certainly simplify the process of translation, they cannot fully replace it.

## Benefits

Conventional translation processes require immense expertise, which narrows down the number of people who translate. The digital methods discussed in this article can make trans-creation a much broader activity, enabling many more people to participate in it. We encouraged around 250 student-teachers to translate English Wikipedia pages and publish it on Kannada Wikipedia. Another 300+ teachers edited KOER (Karnataka Open Educational Resources) content in English and Kannada. In fact, this approach also serves as a professional development activity. The APU translation project was part of a continuous professional development program for our translation team, and gave us excellent opportunities to engage deeply with important texts in education. We used these translations in our work with

Kannada medium government schools. Teachers in Karnataka widely use KOER content.

If teams of teachers having diverse linguistic skills, or those teaching languages, can be encouraged to come together to take on the translation of articles relating to the education domain, it will have two benefits:

- a) it would increase the number of articles available in the target language,
- b) it would support the continuous professional development of teachers.

The scalability of the translation exercise depends on institutionalizing the energies and scaffolding the process. If we see participation as a spectrum, this model moves us further in the direction of democratization of resource-making. In this article, the SL was English. However, a similar approach would work for other SL's as well, as most Indian teachers are multilingual.

IT for Change is an NGO working for a just and equitable digital society, and believes in CCVYSA copyright (Creative Common BY Share and Adapt)

## References

- Bhattacharyya, P. (2004). *Machine translation, language divergence and lexical resources* [PowerPoint slides]. [https://cupdf.com/document/machine-translation-language-divergence-and-lexical-res\\_Hlt57728083\\_Hlt57728084o\\_Hlt57728083\\_Hlt57728084urces-pushpak-bhattacharyya.html](https://cupdf.com/document/machine-translation-language-divergence-and-lexical-res_Hlt57728083_Hlt57728084o_Hlt57728083_Hlt57728084urces-pushpak-bhattacharyya.html)
- Content Translation (n.d). In *Wikipedia*. [https://www.mediawiki.org/wiki/Content\\_translation](https://www.mediawiki.org/wiki/Content_translation)
- English Wikipedia. (n.d.). In *Wikipedia*. [https://en.wikipedia.org/wiki/English\\_Wikipedia](https://en.wikipedia.org/wiki/English_Wikipedia)
- Kannada Wikipedia (n.d.). [https://en.wikipedia.org/wiki/Kannada\\_Wikipedia](https://en.wikipedia.org/wiki/Kannada_Wikipedia)
- Karnataka Open Educational Resources (KOER) project. (n.d). [https://karnatakaeducation.org.in/KOER/en/index.php/Main\\_Page](https://karnatakaeducation.org.in/KOER/en/index.php/Main_Page)
- List of Wikipedias. (n.d.). In *Wikipedia*. [https://en.wikipedia.org/wiki/List\\_of\\_Wikipedias](https://en.wikipedia.org/wiki/List_of_Wikipedias)
- MediaWiki. (n.d.). In *Wikipedia*. <https://www.mediawiki.org/wiki/MediaWiki>
- NCERT.(2013). *Information and communication technology for school system: Curricula for ICT in education. Version 1.01*. Central Institute for Educational Technology, NCERT.
- Papert, S., & Harel, I. (1991). Situating constructionism. In S. Papert & I. Harel (Eds.), *Constructionism* (pp. 1-11). Ablex Publishing. <http://www.papert.org/articles/SituatingConstructionism.html>
- Sarukkai, S. (2013). Translation as method: Implications for history of science. In B. Lightman, G.McQuat & L.Stewart (Eds.), *The circulation of knowledge between Britain, India and China* (pp. 309–329). Brill.
- Shivaprakash H.S. (2020, August 10). *CUK International Webinar* [Webinar]. <https://youtu.be/jNiENOc0jfU>