

तकनीकी अनुवाद एवं अनुसृजन पर पाठ्यक्रम
Course on Technical Translation & Transcreation

Concept and designed by,

Dr Om Vikas, (Min. of Electronic & IT); Late Prof K K Goswami, (Ex-Prof, Delhi);

Prof Nisheeth Joshi, (Banasthali Univ); Prof. Avanish Kumar, (CSTT, New Delhi);

Dr Priyanka Jain, (C-DAC)

dr.omvikas@gmail.com; dravanishkumar@gmail.com;

jnisheeth@banasthali.in, priyankaj@cdac.in

Submitting for approval to,

MoE, AICTE, UGC, DST, AIIMS, New Delhi

Acknowledgment:

Appreciating the team members for giving wings to my dream, I would like to express my humble and sincere regards to each of them for their constant support, patience and self-motivation during the period of the compilation of this course on Technical Translation. This course brought out for every Engineer, Scientist and Professional of innovation and knowledge Creator. My immense appreciation goes to AICTE, AIIMS, UGC, DBT, DST, Department of Higher Education, Ministry of Education, Government of India, Department of Official Languages, Ministry of Home, Government of India, and many more such Institutions, who have motivated and supported me in order to work on Technical Translation and Transcreation.

I am thankful to Late Professor K K Goswami, Gurugram, (Delhi University, Delhi) Professor Avanish Kumar, Bundelkhand University, Jhansi, Professor Nisheeth Joshi (Banasthali University), Dr Priyanka Jain, (C-DAC), Dr Vijay Kumar, (MEITY), and other academic well-wishers who have provided very fruitful suggestions and make it possible in this current form. Further, I would like to thank all academicians, experts, and fellow researchers, who helped in completing this task. With the heartfelt gratitude to all the co-passengers in this journey, I finally came up with this work and profusely thank the invisible hands as well.

Dr. Om Vikas, (Former Director, IITM, Gwalior) New Delhi

Executive Summary

English was adopted as the medium of Instruction in higher studies of Science and Technology. This has resulted in outshining Indian talent at global level. But it retarded scientific and technological communication with people at large and became obstacle in promoting inclusive innovation and entrepreneurship. Innovation is the key for economic growth and improving quality of life of the people at large. Translation is essentially Transfer of Knowledge from one Language into another language. Technical Translation and Transcreation course is designed to come out from this dilemma of English language for the large non English population of India. Direct interaction in the language of society with educated youths with knowledge and skills opens avenues for entrepreneurship and effective solution to the societal and economic problems. Design of a curriculum of three credits for about 50 hours will open

new and advance avenues to enter into the domain of technical education as well medical education and other domain of knowledge. Transcreation अनुसृजन rather than mere translation that demands evaluation on ease of use, communicatively, comprehension and prompting creative ideas. There must be access to dictionaries, parsers, translation memories, term bank, and best practices in scientific writings. Objectives and learning outcomes shall be as under:

- familiarity with some relevant translation software and tools to use them in practical translation
- skills in terminology management and construction of small translation corpora or terminology bank to assist their translation practice

- ability to make use of internet resources for their translation
- research and innovation skills in translation technology and service

The proposed course and curriculum include the **Principles of Translation; Contrastive (SL-TL) Language Analysis: English & Hindi; Language Technology; Translation: Issues, Constraints and Evaluation; MT and practice on Transcreation workbench; and Project.** This has designed with the in-built scope for every knowledge professional and erase the language barrier in the society by and large. It will enhance the fundamental knowledge as well skill and create the job opportunity in translation field.

Course Guidelines on Technical Translation & Transcreation

Contents

1. Preamble
2. Knowledge for Innovation
3. Skill Development for Knowledge Transfer
4. Curriculum for Technical Translation & Transcreation
5. Tentative Curriculum Guidelines

1. Preamble

NEP 2020 suggests promotion of Indian languages in education system from school level to college and university levels. This will enhance greater comprehension of the subject and connect with society. Direct interaction of society with educated youths with knowledge and skills opens avenues for entrepreneurship and effective solution to the societal and economic problems. According to AISHE 2021 report 338 Lakh students enrolled in

UG & PG programs. Maximum enrolment was in Arts & commerce, then in science and engineering. 37.70 lakh students enrolled at the undergraduate level in Engineering. At the PG level enrollment in all branches was 12.6% of total enrolment. PhD students were 203 thousand (2019-20). Maximum numbers of PhDs enrolled were in Science, Engineering & Technology. According to AICTE statistics 2021-2022, there are about 9,000 institutions with intake of about 29.7 lakhs at UG, PG & Diploma levels. According to National Medical Commission, number of medical colleges are 612 with seats of about 92 thousands. All these graduates in science, engineering and medical science will have to interact with society in Indian language for effective service to society. At present the higher education in these disciplines had been in English. There is need to empower them with skills of effective, comprehensible,

easy to understand knowledge transfer from English into Indian language.

Online Machine translation tools are becoming easily available. In order to provide comprehensible translation text of lecture notes / book, it requires editing by the faculty / pass out graduates at workplace. Their direct involvement is essential for continual improvements, and promoting co-creation efforts. Curriculum Guidelines of a 4 credit are suggested below as per the NEP 2020 for the UG / PG level technical Programs running in various HEIs.

2. Knowledge for Innovation

There had been rapid advancement in Science & Technology during 20th century. In 21st century advancements in S&T are even more aggressive. Socio-economic development of a country depends largely on the knowledge available in their people's language. Countries having national language that is non-English have resorted to Translation largely from English to their language, for example from English into Japanese, Korean, etc. Advanced nations also sought S&T knowledge of the competing nations through Translation process from Foreign Language into their Language, for example from Russian into English.

In order to ease the process of Translation, number of efforts of have been made since 1960s for developing Machine Translation systems. In India also during last two decades some projects on Machine Translation between English and Indian Languages, and between Indian Languages to Indian Languages have made good progress. MT alone may not give comprehensible and intelligible translation. But this is like a draft available for quick correction, editing and improvising presentation style to suite to users.

MT is acceptable if 80% or more of human efforts are saved and the rest is left for post editing. Since 1960s, India tried to improve economic growth with focus on Education on Science & Technology. Literacy was to be improved. Several technical institutions were set up in public and private sector. Medium of instruction up to secondary education in most of states was regional language. Hindi was made official language. Hindi grew as a link language between people.

However, English was adopted as the medium of Instruction in higher studies of Science and Technology. This has resulted in outshining Indian talent at global level. But it retarded scientific and technological communication with people at large and became obstacle in promoting inclusive innovation and entrepreneurship. Innovation is the key for economic growth and improving quality of life of the people at large.

3. Skill Development for Knowledge Transfer

In the backdrop of the above, it is proposed to design a curriculum of three credits for about 50 hours that may be submitted to AICTE for approval as an elective credit course under Choice Based Credit System (CBCS) scheme. The same scheme has been approved by UGC as well. This will open avenue to enter into the domain of technical education. Similar efforts may be made for approval as elective credit course in Medical Institutions under IMC.

Important is to use technology to augment creative technical writing. Focus need be on Transcreation अनुसृजन rather than mere translation that demands evaluation on ease of use, communicatively, comprehension and prompting creative ideas. There must be access to dictionaries, parsers, translation

memories, term bank, and best practices in scientific writings.

Translation is essentially Transfer of Knowledge from one Language into another language.

अनुसृजन (Transcreation)

भाषा और संस्कृति भिन्नता के कारण अनुवाद एक तथ्यात्मक एवं सृजनात्मक प्रविधि है। पाठकोन्मुखी सुबोध अनुवाद को अनुसृजन कह सकते हैं।

वर्तमान संदर्भ में अंग्रेजी को स्रोतभाषा से लक्ष्य भाषा हिन्दी में अनुवाद की चर्चा करते हैं।

अनुवाद सामग्री साहित्यिक अथवा साहित्येतर हो सकती है। साहित्यिक के अंतर्गत कथा, कहानी, उपन्यास, कविता, महाकाव्य आदि।

साहित्येतर के अंतर्गत तकनीकी विषय (जैसे विज्ञान, आयुर्विज्ञान, प्रबंधन, प्रौद्योगिकी, विधि, वाणिज्यिक, प्रशासनिक, जन संचार इत्यादि)। इन सबकी अपनी अपनी विशिष्ट शब्दावली हैं, वाक्य पदीय मानक हैं, और सामाजिक जीवन मूल्यों के समावेशन की अपेक्षा है। सुबोध एवं रोचक अभिव्यक्ति के लिए विशिष्ट लेखन शैली है।

प्रौद्योगिकी की गति, प्रबल प्रभावकारिता और प्रसार से विश्व समाज प्रभावित हुए बिना नहीं रह सके। विगत 20वीं सदी से प्रारंभ में 10,000 विश्व भाषाएं जीवित थी, सदी के अंत तक लगभग 6,700 विश्व भाषाएं बच सकीं। प्रति वर्ष 2 प्रतिशत विश्व भाषाओं का लोप होता जा रहा है। भाषा के लोप से लोक संस्कृति और परंपरागत ज्ञान का विलोप होता है। भाषिक आदान-प्रदान तत्काल मौखिक संभव है। लिपि के माध्यम से सुरक्षित रखा जा सकता है। सभ्यता के विकास और आवागमन के बढ़ने से भाषाएं और लिपियां एक दूसरे से प्रभावित हुईं।

पाणिनी जैसे मनीषियों ने ध्वनि एवं लेखन में ऐक्य पर बल देते हुए ध्वनियों का स्वर एवं व्यंजन में वर्गीकरण किया, उच्चारण स्थान और विधि के आधार पर लिपि संरचना सारणी

बनायी। देवनागरी ध्वन्यात्मक है, वैज्ञानिक आधार है। लिपि-व्याकरण भी दिया।

- उच्चारण-स्थान (P) : velar (कंठ्य), palatal (तालव्य), retroflex (मूर्धन्य), dental (दंत्य), labial (ओष्ठ्य)
- उच्चारण की विधि (M) :
voice-less (अघोष)-Unaspirated (अल्पप्राण) अघ-अप
voice-less (अघोष)-Aspirated (महाप्राण) अघ-मप्र
voiced (घोष)-Unaspirated (अल्पप्राण) घ-अप्र
voiced (घोष)-Aspirated (महाप्राण) घ-मप्र
nasal (नासिक्य), semivowel/vowel derivatives, fricative voiceless and voiced

ज्ञानान्तरण / अनुवाद पाथेय

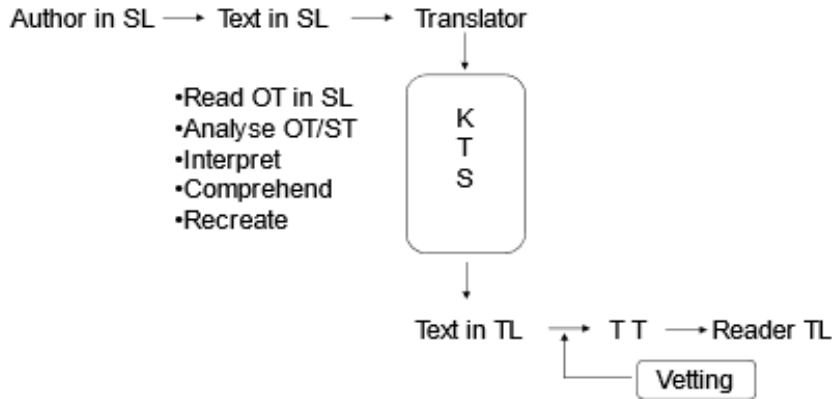
- वक्ता/लेखक → ज्ञानान्तरक → श्रोता/ पाठक (भाषा-i, संस्कृति-j) (भाषा-k, संस्कृति-m)
ज्ञानान्तरक / अनुवादक की जानकारी सीमित होने से भूलों की संभावनाएं अधिक, और सही जानकारी ढूंढने में अधिक समय लगता है। अनुवाद का काम अधिक होने और तत्काल मांग के कारण अनुवाद की गुणवत्ता घटती है। इसलिए मशीन की मदद से अनुवाद का प्रारम्भिक प्रारूप लिया जा सकता है। ज्ञानान्तरण के दो पक्ष हैं – 1. अनुवाद जो लेखकोन्मुखी होता है, और 2. अनुसृजन जो पाठकोन्मुखी होता है। ज्ञानान्तरक से लक्ष्य भाषा और संस्कृति की जानकारी अपेक्षित है।
- भाषा के संबंध में अपेक्षित जानकारी वर्तनी, व्याकरण, लेक्सिकोन, शब्दकोश, संख्या लेखन, दिन-मास-वर्ष, ऋतुएँ, मुहावरा कोश, समांतर कोश, एनोटेटेड कॉर्पस, वाक्यविन्यास, संधि विच्छेद, अलंकार, उत्कृष्ट लेख संग्रह, शैली,

उच्चारण वैशिट्य इत्यादि ... का भाषायी ज्ञान ।

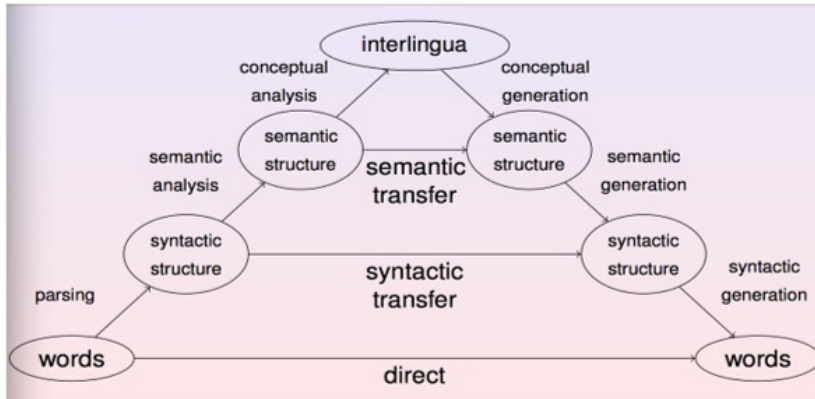
- संस्कृति के संबंध में अपेक्षित जानकारी रीति रिवाज, उत्सव, नृत्य, संगीत, विवाह पद्धति, कार्य पद्धति, भाव भंगिमा, लोक कला ... इत्यादि ।

ज्ञानांतरण / Knowledge Transfer System

(अनुवाद लेखकोन्मुखी होता है, और अनुसूजन पाठकोन्मुखी होता है)



Machine Translation Triangle



Machine Translation Approaches :

प्रविधि आधार – व्याकरण नियम, द्विभाषिक कॉर्पस, प्रयुक्ति उदाहरण, शब्द-पद सांख्यिकी, शब्दस्तर, वाक्य स्तर, अंतर-भाषा

- Rule-Based MT (RBMT) : Rationalism (morphological, syntactic & semantic analysis) more time for MT & less storage requirement
- Corpus-Based MT (CBMT) : Empiricism (use bilingual text corpora) less time for MT & more storage requirement
- Direct, Transfer & Interlingua MT
 - *Direct*: word by word level with simple grammatical adjustments
 - *Transfer*: analysis, syntactic & semantic structural abstraction and generation In both cases SL(n) to TL(k) translation will require $n \times k$ transfer steps,
 - *Interlingua*: analysis into abstract universal representation and generation. SL(n) to TL(k) translation will require $n + k$ transfer steps.
- Statistical (SMT) : based on statistical models of bilingual text corpora

- Example-Based MT (EBMT) : based on parallel corpora and 3 steps of example acquisition, example base management and example application & synthesis

MT Systems developed with funding by Min. of Electronics & IT, Govt of India for Indian Languages:

अंग्रेजी से हिन्दी में अनुवाद प्रणालियाँ -

- Anusarak, Matra, Mantra, Angla-Bharti

बहु भाषिक अनुवाद प्रणालियाँ -

- SAMPARARK (IL-IL) P-H, H-P, U-H, Te-Ta
- ANUVADAKSH (E-IL) IL: H, M, B, O, Ta,U
- ANGLA-BHARTI (E-IL) IL: B, P, Mal, U
- CILA : A, B, G, H, M, O, P, Ta, Te

IL: Indian Language, E: English, A: Assamese, B: Bengali, G: Gujarati, H: Hindi, K: Kannada, M: Malayalam, O: Oriya, P: Punjabi, Ta: Tamil, Te: Telegu, U: Urdu

Online Translators (in the context of Indian Languages)

- Google, Microsoft, Facebook, Yahoo
- www.Translate.google.com/manager
- www.Bing.com/translator

Other Tools needed (to develop) -

Cloud based Transcreation workbench including Spell Checkers, Grammar Checkers, Terminology Manager, E-dictionaries, Translation Memory Tools, Localization tools, etc.

मशीनी अनुवाद की व्यापारिक संभावनाएँ हैं - T2S (टेक्स्ट से बोल), S2T (बोल से टेक्स्ट), S2S (बोल से बोल)

मूल्यांकन - Evaluation of Machine Translation

- A metric that evaluates machine translation output represents the quality of the output. The measure of evaluation for metrics is correlation with human judgment.
- **BLEU** was one of the first metrics to report high correlation with human judgments of quality.
- The **METEOR** metric is designed to address some of the deficiencies inherent in the BLEU metric. The metric is based on the weighted harmonic mean of unigram precision and unigram recall.
- **LEPOR** yielded higher system-level correlation with human judgments than several existing metrics such as BLEU, Meteor-1.3, TER, AMBER and MP4IBM1.
- **COMET** is a robust Crosslingual Optimized Metrics for Evaluation of Translations by ensembling systems that model different aspects of MT evaluation.
- **MATESE** assign a label to each token of the candidate. These labels identify error spans, together with their severity, chosen among Major and Minor. Finally, in order to associate a score with the entire tagged sentence

अनुवाद गुणवत्ता का मूल्यांकन

- भाषा-शैली - शुद्धता, औचित्य, सहजता, सुसंगतता, शैली
- शब्दावली-संदर्भ-मानव मूल्य- फार्मूला, नाम पद, प्रक्रिया पद, मानव मूल्य, संदर्भानुकूल
- त्रुटि-विश्लेषण स्रोत भाषा में - वर्तनी, व्याकरण, शब्द चयन, अर्थ, शैली
- त्रुटि-विश्लेषण लक्ष्य भाषा में - जोड़ा, हटाया, अर्थ-विकृति, अतिरिक्त भाषा संदर्भ, भाषिक

भिन्नता

- पाँच माप बिन्दु : 5— उच्चतम, 4— उच्च, 3— सामान्य, 2— न्यून, 1— न्यूनतम
- गुण = $\sum ((क1, क2, क3, क4, क5) + (ख1, ख2, ख3, ख4, ख5))$
- दोष = $\sum ((ग1, ग 2, ग 3, ग4, ग5) + (घ1, घ2, घ3, घ4, घ5))$
- सकल गुण अधिकतम, और सकल दोष न्यूनतम किए जाएं।

अनुसृजन क्षेत्र विशाल है – कार्य और शोध की दृष्टि से। 2005 में गठित राष्ट्रीय ज्ञान आयोग ने लोक भाषा में विविध प्रकार के ज्ञान की कमी को भरपाई के लिए पांच प्रमुख सिफारिशों की :

- सुगमता – ज्ञान प्रप्ति की सुगमता हो
- संकल्पना प्रधानता – शिक्षा के सभी स्तरों पर कंसेप्ट (संकल्पनाओं), उनके अन्तर संबंधों, कारण–प्रभाव विश्लेषण और नवाचार पर बल दिया जाए।
- ज्ञान सर्जना – सामाजिक एवं आर्थिक दृष्टि से उपयोगी ज्ञान सर्जना को बढ़ाया जाए (विशेषतः विज्ञान एवं प्रौद्योगिकी, बौद्धिक संपदा अधिकार, उद्यमिता को)
- अनुप्रयोग (Applications) – ज्ञान तंत्र/नॉलेज सिस्टम का विकास किया जाए।
- सेवाएं – ई–गवर्नेंस जैसी जन–सेवाएं।

राष्ट्रीय ज्ञान आयोग ने राष्ट्रीय अनुवाद मिशन प्रारंभ करने की भी सिफारिश की। कार्य बिन्दु थे : अनुवाद शिक्षण, सूचना प्रसारण, उत्कृष्ट कोटि के अनुवाद को प्रोत्साहन, मशीन अनुवाद व्यवहार में।

2022 में भारत सरकार ने नेशनल ट्रांसलेशन मिशन (NTM) के लिए 500 करोड़ रुपए का प्रावधान किया है।

ज्ञातव्य है कि अनुवाद लेखकोन्मुखी होता है, और अनुसृजन पाठकोन्मुखी होता है।

अपेक्षा है कि इनकी उपलब्धता और प्रयोग–प्रसार को सुनिश्चित किया जाए :

- इनपुट – मानक INSCRIPT की–बोर्ड, मोबाइल डिवाइस पर भी हों,
- कोडिंग – ISCII, UNICODE, Phonicode & Indian Phonetic Alphabet (to develop)
- डिस्प्ले एवं प्रिंटिंग के लिए फॉन्ट एवं फॉन्ट परिवर्तक
- लिप्यंतरण, मशीन अनुवाद, एप्लीकेशन SW
- अक्षर पहचान – OCR, Handwriting CR
- टेक्स्ट टू स्पीच, स्पीच टू टेक्स्ट, स्पीच टू स्पीच
- वेब ब्राउसर, डोमेन नेम, (डॉट) भारत (like .in, .cn, .jp, ...)
- नागरी में यूटिलिटी सॉफ्टवेयर – लाइब्रेरी, एकाउंटिंग, स्कूल–कॉलेज, प्रबंधन, ट्रांसपोर्टेशन, पाठ–लेखन, ऑथरिंग आदि सॉफ्टवेयर नागरी लिपि में मुफ्त और मुक्त सर्वसुगम हों।
- नागरी लिपि–व्याकरण मानकीकरण भाषा सापेक्ष होगा, हिन्दी, मराठी के लिए अलग होगा। नागरी में कंटेंट क्रियेशन और Web Service इंटीग्रेशन और XML आदि के मानकों पर भी काम किया जाए।

मानकीकरण

- नागरी लिपि–व्याकरण मानकीकरण भाषा सापेक्ष होगा। हिन्दी और मराठी के लिए अलग हो सकता है।
- मानक लेखन के सुझाव हों, नए प्रयोगों के प्रति सहिष्णुता रहे। (किया, किये/किए, नयी/नई,)
- नागरी में कंटेंट क्रियेशन और Web Service इंटीग्रेशन और XML आदि के मानकों पर भी काम किया जाए।
- ध्वनि इकाई (unit) का माप–मानकीकरण हो।
- संस्कृत एवं अन्य भारतीय भाषाओं में लिप्यंतरण मानकीकरण

- नागरी टूल्स और टैक्नोलोजी के मानकीकरण और प्रभावी अनुपालन की आवश्यकता है।

नागरी-हिन्दी में प्रयोग संभावनाएं

- Glyph based encoding standard (one byte) ASCII for Roman script and ISCII for Roman + an Indic script. Now (two byte) standard UNICODE for over 100 languages.
- Localization लोकीकरण : BOSS, Firefox, Thunderbird (email SW), ECKO (Content Management)
- 7Bn people + 1018 (zeta)byte content = Big data disconnect ?
- Language Switch: Browser-based Localization, Localization of databases, eg. half past ten साढ़े दस
- Homophone Engine for searches, e.g. Chowdhary – 64 variants, Upadhyay – 28 variants, हिन्दी, हिंदी
- Blogs, Social Networking, e-book
- Free Hindi fonts and software downloadable from <http://ildc.in> -
- Fonts, KB drivers, Open Office, Browser, typing tutor, code/font converters, lexical analyzer, POS tagger, TTS, OCR, Transliteration

लोकीकरण (L10N) एवं वैश्वीकरण (I18N)

- प्रौद्योगिक उत्पाद एवं सेवाओं के अनुवाद में : अन्य भाषा से अपनी भाषा के अनुवाद के समय लोकीकरण (L10N : Localization), और अपनी भाषा से अन्य भाषाओं में अनुवाद के संदर्भ में वैश्वीकरण (I18N : Internationalization) की आवश्यकताओं को ध्यान में रखना पड़ता है।
- कार्य क्षेत्र विषयक सुझाव — Localization of MOOCs (Massive Open Online Courses) and other e-learning resources at school and college level. Localization of

vocational / skill training material using open localization tools.

- नीति विषयक सुझाव - Mandatory Introduction of Indian language processing in IT courses at secondary, senior secondary and tertiary levels – UG & PG. Incentives for SOHO knowledge (multilingual) industry, and Quality Certification.
- प्रौद्योगिकी विषयक सुझाव - “Localization Cloud” to automatically download the requisite tools. Development of Universal Voice based Communication Code (PHONICODE) and other requisite technologies - Spell Checkers, Grammar Checkers, Terminology Manager, E-dictionaries, Translation Memory Tools, etc.

4. Curriculum for Technical Translation & Transcreation

4.1 Objectives and learning outcomes

At the end of the course, a student should be able to demonstrate . . .

- familiarity with some relevant translation software and tools to use them in practical translation
- skills in terminology management and construction of small translation corpora or terminology bank to assist their translation practice
- ability to make use of internet resources for their translation
- research and innovation skills in translation technology and service

4.2 Workload

This course will be taught over 15-16 weeks with 3 hours classroom contact per week. There will be focus on case studies and practical sessions. Project work will also be there at end of the course.

4.3 Scope and syllabus

This course is to train students to apply their theoretical and conceptual background to a very practical approach to translation using machine translation (MT), parallel corpora, translation memory (TM) management, terminology database (TD) management, translation project management, translation of documentation, and software localization to exploiting translation resources available on the internet and legacy translation data.

The course will teach students to make effective use of specialized tools to enhance their translation productivity, share data and manage projects. It will allow students to develop not only practical expertise but also a critical faculty for evaluating their relative merits. The students will have access to applications widely used in the language industries for TM management, TD management, project management, software localization and subtitle translation. The course will focus on a critical appraisal of the usefulness and usability of the tools and students are encouraged to contribute to the course by bringing to the class their own surveys and investigations into the use of tools relevant to translation technology.

Skills developed in translation technology, translation theory and translation practice will enable students to improve both their translation efficiency and translation quality, and hence considerably increase their competitiveness. As spin off advantage of this course, a student may develop skill in original creative technical writing in Hindi / Indian Language. This may facilitate the Government mission on AWSAR (Augmenting Writing Skill in Articulating Research). The resultant piece of research may benefit entrepreneurs to improvise products and services, and thus scaling up Inclusive Innovation.

4.4 Choice of Course

- College and skill partner can jointly issue a

certificate to the student additionally.

- Credit distribution:
 - 1 credit (theory) = 15 hours
 - 1 credit (training) = 30hours

4.5 EVALUATION

The performance of a student in each course is evaluated in terms of percentage of marks with a provision for conversion to grade point. The evaluation may be continuous and holistic based on the following assessment parameters: Academic, Technical Skill, Behavioral Skill, and Extra-curricular.

4.6 Grades / Marks in percent

O: (90-100%), A: (75-90%), B: (60-75%), C: (40-60%), F: (<40%)

Marks obtained in decimals will be converted in nearest integer.

4.7 Topics to be covered include:

- 1.Principles of translation and Critical review of translation technology
- 2.Introduction to MT and online MT tools as an aid to translators
- 3.Introduction to Translation memory (TM) & TM management, Terminology database (TD) & TD management
- 4.Translation projects management
- 5.Corpora (monolingual, parallel bilingual and comparable) as an aid to translators
- 6.Principles and skills in localization (software and websites), and subtitle creation
- 7.Practice and Critical evaluation of Transcreation workbench (MT and CAT tools)

5. Tentative Curriculum Guidelines

Unit 1. Principles of Translation

Unit 2. Contrastive (SL-TL) Language Analysis: English & Hindi

Unit 3. Language Technology

Unit 4. MT and practice on Transcreation workbench (with Semantic, Discourse and Anaphora etc)

Unit 5. Evaluation of Translation: issues and limitations

Unit 6. Project

There is a valid comment by Prof. Avanih Kumar. Accordingly there is correction in the table on tentative distribution of marks. class, practical and projects as shown below.

Tentative time distribution of Theory (60%), Practical (40%), & marks 80% for Theory and practical assessment and 20% marks for Project work that may be in grade A or B. This may correspond to 20% marks. We may not specify the number of hours needed. We need not specify the time for Project.

UNIT	DESCRIPTION	Theory (60%)	Practical (60%)	Marks 80% + Grade
UNIT-1	Principles of Translation	12	—	12
UNIT-2	Contrastive (SL-TL) Language Analysis: English & Hindi	12	8	16
UNIT-3	Language Technology	12	12	18
UNIT-4	Translation: Issues, Constraints and Evaluation	12	8	16
UNIT-5	MT and practice on Transcreation workbench	12	12	18
UNIT-6	Project	—	—	Grade A/B

UNIT-1: अनुवाद: सिद्धांत पक्ष (Principles of Translation)

क. अर्थ और अवधारणा – परिभाषा, स्वरूप, प्रकृति और प्रक्रिया

ख. अनुवाद के प्रकार

1. पूर्ण- आंशिक
2. समग्र और परिसीमित
3. साहित्यिक (सर्जनात्मक, भावानुवाद)
4. शाब्दिक अनुवाद
5. सार अनुवाद एवं छायानुवाद
6. यांत्रिक अनुवाद (मशीनी)

ग. 1. संरचना पक्ष (ध्वनि पक्ष, लेखिम पक्ष, शब्दपक्ष, वाक्यपक्ष)

2. शब्दाकेश
3. अनुप्रयुक्त पक्ष एवं प्रोक्ति

घ. अनुवाद की उपादेयता

1. भाषा विकास के संदर्भ में
2. द्विभाषिक एवं बहुभाषिकता के संदर्भ में
3. राजभाषा, उद्योग एवं तकनीकी विकास के संदर्भ में
4. भारत की भावात्मक एकता के संदर्भ में अनुवाद चिन्तन का क्रमिक विकास: भारतीय और पाश्चात्य संदर्भ

अनुवाद समतुल्यता का सिद्धांत

अनुवादक की कुशलताएँ और दक्षताएँ, ज्ञान तथा गुण

सहायक सामग्री: कोश, शैली सहायिका, कंप्यूटर ज्ञान

संदर्भ पाठ्य पुस्तकें -

1. अनुवाद कला: कुछ विचार – डॉ. आनंद प्रकाश खेमाणी, प्र. एस. चांद एण्ड कम्पनी, दिल्ली
2. अनुवाद विज्ञान – डॉ. भोलानाथ तिवारी,

प्र. शब्दकार, दिल्ली

3. अनुवाद कला – चारुदेव शास्त्री
4. अनुवाद – अवधारणा और अनुप्रयोग – (सं.)
डॉ. चंद्रभान रावत, डॉ. दिलीप सिंह, द.भा.हि.प्र.
सभा, मद्रास
5. अनुवाद: मूल्य और मूल्यांकन – डॉ. शशिमुदिराज
6. Nida, E.A. – Language Structure and Translation, Stanford University Press.
7. Nida and Taper – The Theory and Practice of Translation, Stanford University Press.
8. Catford, C.J. 1966 – Linguistic Theory of Translation, OUP, Navjeevan Publishing House, New Delhi.
9. T. Savory – Art of Translation, London: Cape.

UNIT-2: व्यतिरेकी विश्लेषण और पाठ विश्लेषण (Contrastive (SL-TL) Language Analysis: English & Hindi)

व्यतिरेकी भाषाविज्ञान (सिद्धांत एवं व्यवहार)

1. तुलानात्मक भाषाविज्ञान और व्यतिरेकी भाषा विज्ञान: परिभाषा, स्वरूप और प्रविधि
2. संरचनात्मक भाषाविज्ञान में भाषा तुलना का महत्त्व
3. व्यतिरेकी विश्लेषण की प्रक्रिया
4. हिंदी-अंग्रेजी में वाच्यीकरण
5. हिंदी-अंग्रेजी में क्रिया संरचना की तुलना
6. हिंदी-अंग्रेजी में प्रेरणार्थीकरण की तुलना
7. हिंदी-अंग्रेजी में निजवाचकता
8. हिंदी-अंग्रेजी में क्रिया विशेषण पदबन्ध
9. हिंदी-अंग्रेजी में सम्बन्ध वाचक वाक्य
10. हिंदी की विशिष्ट संरचनाओं का अनुवाद: रंजक क्रियाएँ, निरन्तरता बोधक क्रियाएँ, 'ने' संरचना, सर्वनाम, सम्प्रदान में कर्ता।
11. अंग्रेजी की विशिष्ट संरचनाओं का अनुवाद: have, उपसर्ग और उपसर्गीय प्रयोग।

पाठ विश्लेषण

1. भाषा और सम्प्रेषण: भाषा संरचना की इकाई वाक्य, सम्प्रेषण की इकाई पाठ, अर्थन्विति (शब्दार्थ, वाक्यार्थ, पाठार्थ), सन्दर्भ
2. पाठ, प्रोक्ति, महावाक्य: पाश्चात्य और भारतीय सन्दर्भ, परिभाषा स्वरूप, प्रोक्ति के मुख्य अंग: (संसक्ति और अर्थ संगति)
3. पाठ संरचना (Text Structure): वाक्य संरचना स्तरों का अधिक्रम अन्तर वाक्य योजक (Cohesive Ties) पाठीय अन्विति
4. पाठ, प्रोक्ति और अनुवाद: अन्तःवस्तु और शैली।

सहायक ग्रंथ

1. अनुवाद व्याकरण: सूरज भान सिंह, प्रभात प्रकाशन, दिल्ली
2. व्यतिरेकी भाषाविज्ञान-विज्ञय राघव रेड्डी
3. अंग्रेजी-हिंदी वाक्य संरचना का व्यतिरेकी अध्ययन- रमेशचंद्र शर्मा
4. अनुवाद विज्ञान की भूमिका: कृष्ण कुमार गोस्वामी, राजकमल प्रकाशन, दिल्ली
5. व्यतिरेकी भाषाविज्ञान: भोलनाथ तिवारी, सं. आलेख प्रकाशन, दिल्ली.
6. अनुप्रयुक्त भाषा विज्ञान: रवीन्द्रनाथ श्रीवास्तव, भोलानाथ तिवारी, कृष्ण कुमार गोस्वामी (सं.) आलेख प्रकाशन, दिल्ली
7. पाठ विश्लेषण: प्रो. दिलीप सिंह, वाणी प्रकाशन, दिल्ली
8. Harris, Z. S., & Harris, Z. S. (1970). Discourse analysis (pp. 313-348). Springer Netherlands.
9. Kumar, S. (2003). Stylistics and Language Teaching: With a Section on Translation Studies. Kalinga Publications.

UNIT-3: भाषा प्रौद्योगिकी (Language Technology)

1. आधुनिक भाषा विज्ञान और कंप्यूटरीकृत

भाषाविज्ञान

2. भाषा प्रौद्योगिकी: अर्थ, स्वरूप और क्षेत्र
3. भाषा प्रौद्योगिकी में आधुनिक भाषा विज्ञान का अनुप्रयोग—भाषिक और भाषा अभियांत्रिकी का संदर्भ
4. सूचना प्रौद्योगिकी और हिंदी भाषा
5. प्राकृतिक भाषा संसाधन – (NLP)
6. पार्सिंग: परिभाषा एवं प्रकार – टाप डारुन और बाटम अप
7. ज्ञान प्रतिपादन तथा उसकी विभिन्न तकनीकें
8. कार्पस निर्माण: सिद्धांत, प्रकार, बी.एन.सी. भारत में कार्पस निर्माण में पाणिनी टैग सेट।
9. कार्पस निर्माण का उपयोग – वर्तनी शोधक, व्याकरण शोधक, शैली शोधक, प्रकाशिक वर्ण (OCR)
10. कंप्यूटर के लिए पदसमूह (Phrases), व्याकरण, वर्डनेट, सिमेन्टिक नेट का निर्माण।
11. वाक् संश्लेषण और वाक् विश्लेषण स्वरूप और व्यावहारिक समस्याएँ।
12. वाक् से पाठ और पाठ से वाक्

सहायक ग्रंथ

1. Allen, J. (1995). Natural language understanding. Benjamin-Cummings Publishing Co., Inc.
2. Rich, E., & Knight, K., Rao, S.B., Artificial Intelligence McGraw-Hill. New York.
3. Winograd, T. (1983). Language as a cognitive process: Volume 1: Syntax., Stanford University
4. Grishman, R. (1986). Computational linguistics: an introduction. Cambridge University Press.
5. Applied Linguistics: R.N. Shrivastav, Kalinga Publications, Delhi
6. Artificial Intelligence: Andres, A.M.

(Abacus Press).

7. सूचना प्रौद्योगिकी विशेषांक: गवेषणा और अनुवाद पत्रिकाएँ
8. Bharti, A., Chaitanya, V., Sangal, R. Natural Language Processing: A Paninian Perspective, Prentice Hall of India, New Delhi.
9. Natural Language Processing, Harry, Tennans, Pertro celli Books, New York.

UNIT-4: अनुवाद: समस्याएँ, सीमाएँ और मूल्यांकन (Translation: Issues, Constraints and Evaluation)

खंड-क: स्रोत भाषा और लक्ष्य भाषा

1. स्रोत और लक्ष्य—भाषा की भिन्नाताजन्य समस्याएँ
 - अ. भाषिक स्तर पर – संरचनागत और शैलीगत
 - ब. प्रोक्तिपरक
 - स. सांस्कृतिक शब्दों तथा अभिव्यक्तियों के स्तर पर
 1. सांस्कृतिक शब्दावली
 2. मिथकीय अभिव्यक्तियाँ
 3. मुहावरे और लोकोक्तियाँ
2. कविता: बिम्ब, प्रतीक, अलंकार, छन्द, लय, अर्थ और विचार
3. नाटक: संवाद, समाज भाषिक, रंगमंच निर्देश एवं शैली
4. कथा साहित्य: आंचलिक भाषा, परिवेश, कथन—भंगिमा।
5. साहित्येतर अनुवाद की समस्याएँ
 - क. पारिभाषिक शब्दावली
 - ख. संरचना
6. अनुवाद की सीमाएँ
 - (क) भाषिक
 - (ख) सामाजिक—सांस्कृतिक
 - (ग) प्रयुक्तिपरक

7. अनुवादनीयता की समस्याएँ
क. अनुवादनीयता, अनूद्यता की अवधारणा
ख. पूर्ण, आंशिक और शून्य अनूद्यता
- खंड-खः अनुवाद पुनरीक्षण मूल्यांकनः**
 1. अनुवाद पुनरीक्षण के मुख्य मुद्दे: मूलनिष्ठता, बोधगम्यता
 2. अनुवाद पुनरीक्षण की मुख्य विधियाँ : अनुवाद मूल्यांकन और समीक्षा
 3. अनुवाद मूल्यांकन: प्रकृति, उपादेयता, उद्देश्य, प्रविधि एवं सामान्यीकृत निष्कर्ष।
 4. अनुवाद दोष विश्लेषण (मूल्यांकन)
 5. पाठ विश्लेषण और अनुवाद मीमांसा
- सहायक ग्रंथ**
 - 1 अनुवाद की व्यावहारिक समस्याएँ: भोलानाथ तिवारी, महेन्द्र चतुर्वेदी
 - 2 अनुवाद सिद्धांत और समस्याएँ: (सं.) रवीन्द्रनाथ श्रीवास्तव; कृष्ण कुमार गोस्वामी
 - 3 अनुवाद विज्ञान की भूमिका: कृष्ण कुमार गोस्वामी राजकमल प्रकाशन नई दिल्ली
 - 4 विकासशील देशों में अनुवाद की समस्याएँ: बालकृष्ण केलकर (सम्पा.) नेशनल बुक ट्रस्ट नई दिल्ली।
 - 5 अनुवाद सिद्धांत की रूपरेखा: सुरेश कुमार वाणी प्रकाशन नई दिल्ली
 - 6 Translation & Translating: Theory and Practice: R. Bell. London: Longman.
 - 7 Linguistic And Cultural Problems of Translation: Trivedi H.C. New order Book Depot Ahmedabad
- UNIT-5: मशीनी अनुवाद एवं अनुसृजनिका (MT & Practice on Trans-creation Workbench)**
 1. मशीनी अनुवाद: परिभाषा स्वरूप, इतिहास और उपादेयता
 2. सामान्य अनुवाद और मशीनी अनुवाद
(क) मानव साधित (ख) मशीन साधित
 3. मशीन अनुवाद प्रक्रिया: पारसिंग, अन्तरण, जेनेरेशन, पूर्व संपादन, पश्च संपादन, अंतरकोश निर्माण, शब्दस्तरीय कोश, बहुशाब्दिक अभिव्यक्ति कोश।
 4. अनुवाद प्रविधियाँ: प्रत्यक्षण, अंतरण, अंतर भाषा और मशीन अनुशिक्षण
 5. अनुवाद विचलन की जटिलता (Complexity of Translation Divergence)
 6. मशीन अनुवाद और व्याकरणिक विश्लेषण: शब्द कोशीय स्तर पर, व्याकरणिक संरचना के स्तर पर, ट्री एडजाइनिंग ग्रामर, पाणिनीय व्याकरण के स्तर पर।
 7. भारत में मशीन अनुवाद का इतिहास
(क) भारतीय भाषाओं में परस्पर अनुवाद: तेलुगु, कन्नड़, मराठी, बंगला, पंजाबी से हिंदी अनुसारक
(ख) अंग्रेजी से हिंदी अनुवाद तंत्र: मैट, मात्रा, मंत्रा, लीला, शक्ति का सामान्य परिचय।
 8. इंटरनेट और मशीन अनुवाद – Google Translate, Bing Translate.
 9. Evaluation of Translation: issues and limitations (may include Evaluation metrics of MT like BLEU, METEOR, LEPOR, F-Measure, WER (Word Error Rate), COMET, etc. for the intelligibility, fidelity, fluency, adequacy, comprehension, and informativeness.
 10. अनुसृजनिका एवं अभ्यास (Practice on Transcreation Workbench). Transcreation is the Computer Assisted Human Translation for high productivity, high quality intelligible translation with style and ethical relevance. This is facilitated by a Transcreation Workbench (अनुसृजनिका) that may be cloud based with CAT tools like MateCat, LingoHub,

OmegaT, CASMACAT; which use Translation Memory, Terminology bank, parsers, Spell checker, Grammar checker, post-editing tools, Localization tools, and the parallel corpora, dictionaries, and best practices in scientific writings.

सहायक ग्रंथ

1. King, M. (1987). A tutorial on machine translation. Institut pour les études sémantiques et cognitives/Université de Genève.
2. Grishman, R. (1986). Computational linguistics: an introduction. Cambridge University Press.
3. Locke, W. N., & Booth, A. D. (1956). Machine translation of languages. American Documentation (pre-1986), 7(2), 135.
4. Slocum, J. (Ed.). (1988). Machine translation systems (pp. 1-47). Cambridge: Cambridge University Press.
5. Machine Translation: Theoretical & Methodological Issues, Cambridge Uni. Press
6. Mason, S. B. (2019). Translation quality assessment: from principles to practice Machine Translation: Technologies and Applications, Volume 1.
7. Fomicheva, M., & Specia, L. (2019). Taking MT evaluation metrics to extremes: Beyond correlation with human judgments. Computational Linguistics, 45(3), 515-558.
8. Olive, J., Christianson, C., & McCary, J. (Eds.). (2011). Handbook of natural language processing and machine translation: DARPA global autonomous language exploitation. Springer Science & Business Media.
9. Trujillo, A. (1999). Translation engines: techniques for machine translation. Springer Science & Business Media.
10. Koehn, P. (2009). Statistical machine translation. Cambridge University Press.
11. Koehn, P., & Knowles, R. (2017). Six challenges for neural machine translation. arXiv preprint arXiv:1706.03872.
12. Comelles, E. (2022). Machine Translation and the Assessment of Translation Quality. Atlantis. Journal of the Spanish Association for Anglo-American Studies, 233-244.
13. Koehn, P., Barrault, L., Bojar, O., Bougares, F., Chatterjee, R., Costa-jussà, M. R., ... & Zampieri, M. (2022, December). Proceedings of the Seventh Conference on Machine Translation (WMT). In Proceedings of the Seventh Conference on Machine Translation (WMT).