The Study of Religion and History
Volume 3 issue 1 (2025)
www.srhjournal.com
ISSN(Print): 3006-3329 ISSN(Online): 3006-3337

"مشینی ترجمہ کاری کا ارتقا: نظریاتی بنیادیں، عملی اطلاقات اور عصری رجحانات"

*بشریٰ بلال
اسکالر ایم فل اردو، شعبہ اُردو بہاء الدین زکریا یونیورسٹی، ملتان
Bushrabilal585@gmail.com

سید شاہ زیب اقبال
اسکالر ایم فل اردو، غازی یونیورسٹی، ڈیرہ غازی خان
syedshahzabiqbal247@gmail.com

منیر احمد
اسکالر پی۔ایچ۔ڈی، شعبہ اُردو بہاء الدین زکریا یونیورسٹی، ملتان
Munirahmad31000@gmail.com

فاطمہ کاظم
اسکالر ایم فل اردو، شعبہ اُردو بہاء الدین زکریا یونیورسٹی، ملتان
Fatimakazim788@gmail.com

Abstract:

Machine Translation (MT) is a linguistic-technological phenomenon that emerged in the mid-20th century as a result of the convergence between computer science and linguistics. It aims to automatically convert one human language into another through computational processes. This paper explores the theoretical foundations, historical evolution, technological models, and contemporary applications of MT, with a particular focus on under-resourced languages like Urdu. From the philosophical visions of a universal language in the 17th century to the advent of neural machine translation systems like DeepL, Microsoft Translator, and Amazon Translate, this study traces the intellectual journey and technological advancements in the field. The paper examines key MT models including Rule-Based (RBMT), Corpus-Based (CBMT), Statistical (SMT), Neural (NMT), and Hybrid Machine Translation (HMT), emphasizing their strengths and limitations. It also investigates the challenges of applying MT to Urdu due to limited corpora, lack of annotated data, and cultural-

linguistic complexities. Recent research on models like Fully Convolutional Neural Networks (FConv-NN), Long Short-Term Memory (LSTM), and Example-Based MT for English-Urdu translation is critically analyzed. The paper concludes by outlining future prospects and research directions for enhancing machine translation quality in low-resource languages.

Keywords: *Machine Translation, Neural Machine Translation, Urdu-English Translation, Corpus-Based MT, Rule-Based MT, Natural Language Processing, Artificial Intelligence, FConv-NN, LSTM, Hybrid Translation Models.*

انسانی زبان نہ صرف خیالات کے اظہار کا ذریعہ ہے، بلکہ تہذیب، ثقافت، علم اور رشتوں کے باہمی ربط کا سب سے اہم وسیلہ بھی ہے۔ زبانوں کی رنگا رنگی نے جہاں انسانی تہذیب کو وسعت عطا کی، وہیں مختلف لسانی گروہوں کے مابین باہمی رابطے میں پیچیدگی بھی پیدا کی۔ اس تناظر میں ترجمہ کی ضرورت ہمیشہ سے موجود رہی ہے، لیکن بیسویں صدی میں کمپیوٹر سائنس اور لسانیات کے ملاپ نے ایک انوکھا میدان پیدا کیا، جسے ہم مشینی ترجمہ کاری (Machine Translation - MT) کے نام سے جانتے ہیں۔

مشینی ترجمہ کاری ایک ایسی ٹیکنالوجی ہے جو قدرتی زبان کی عمل کاری (Natural Language Processing - NLP) کے تحت ایک انسانی زبان کو دوسری انسانی زبان میں خودکار طور پر تبدیل کرتی ہے۔ یہ عمل محض لغوی ترجمہ پر مشتمل نہیں ہوتا، بلکہ یہ سیاق، نحو، معنیات اور لسانی اسلوب کو بھی پیشِ نظر رکھتا ہے۔ یوں سمجھا جاسکتا ہے کہ مشینی ترجمہ کاری لسانی، منطقی، سائنسی اور سافٹ ویئر انجینئرنگ کا ایک مشترکہ مظہر ہے۔

سترہویں صدی میں عالمی زبان کی تشکیل کا تصور، ارسطو کے لاجک پر مبنی جوابی مشین، اور جان ولکنز کی زبانوں کو اشارتی اشکال میں بیان کرنے کی کوششیں اسی خواب کی ابتدائی علامات تھیں، جس کی تعبیر بیسویں صدی کے وسط میں کمپیوٹر کی مدد سے ہونا شروع ہوئی۔

جیسا کہ یورک ولکس نے لکھا:

> "Machine translation is the first, the boldest, and perhaps the most visionary attempt to mechanize a core human intellectual function — the act of understanding and generating language."
> [1]

یہی وہ علمی دائرہ ہے جس پر یہ مضمون روشنی ڈالے گا۔

مشینی ترجمہ کاری کا خیال نیا نہیں۔ انسانی تاریخ میں مختلف ادوار میں عالمی زبان کی تشکیل کی خواہش بارہا ظاہر ہوئی۔ رینے ڈیکارٹ، لائبنز اور جان ولکنز جیسے مفکرین نے ایسی زبان کا تصور دیا جو منطق اور عالمی اصولوں پر استوار ہو اور جسے دنیا کا ہر انسان سیکھ سکے۔ جان ولکنز نے اپنے مشہور مضمون *"Essay Towards a Real Character and Philosophical Language" (1668)* میں لکھا:

> "A language in which all things and notions might be expressed with absolute clarity, and without ambiguity..."
> [2]

اس سے اندازہ ہوتا ہے کہ مشینی ترجمہ کاری کا خواب، انسانی فکر کی ایک مسلسل جستجو کا نتیجہ ہے۔ تاہم عملی بنیاد پر اس خواب نے بیسویں صدی کے وسط میں شکل اختیار کی۔

مشینی ترجمہ کی پہلی مربوط کوشش 1933ء میں جارج آرتسرونی اور پیوٹر ٹرویانسکی نے کی، جنہوں نے لغت، گرامر اور ترجمہ ایڈیٹر پر مشتمل ایک سہ رخی نظام پیش کیا۔ ٹرویانسکی کے مطابق:

> "ترجمے کے لیے ایک میکانکی نظام ایسا ہونا چاہیے جو پہلے زبان کو منطقی سطح پر پرکھے، پھر اس کے اجزائے ترکیبی کو ترتیب دے، اور آخر میں انسانی تدوین کی جائے۔[3]"

یہ وہ خاکہ تھا جسے بعد میں مشینی ترجمہ کاری کے تین مراحل کی بنیاد سمجھا گیا:

1. ماخذ زبان کا تجزیہ
2. منتقلی (transfer) یا درمیانی نمائندگی
3. ہدف زبان میں ترکیب

مشینی ترجمہ کاری کا عملی آغاز 1954ء میں ہوا جب Georgetown University اور IBM نے ایک محدود نظام تیار کیا، جس نے 250 روسی الفاظ پر مبنی جملوں کو انگریزی میں ترجمہ کیا۔ یہ دنیا کا پہلا public MT demonstration تھا۔ تاہم اس کامیابی کے باوجود، 1966 میں ALPAC رپورٹ نے مایوس کن تصویر پیش کی:

"مشینی ترجمہ کاری نہ تو درست ہے، نہ تیز، اور نہ سستی۔ اس میں انسان کا کردار زیادہ اہم ہے۔"[4]

اس رپورٹ نے امریکہ میں MT کی فنڈنگ تقریباً بند کر دی، اور یہ میدان یورپ، کینیڈا، اور جاپان کی طرف منتقل ہو گیا۔

جوں جوں مشینی ترجمہ کاری نے ترقی کی، اس کے مختلف فنی اور عملی ماڈلز سامنے آئے۔ ان ماڈلز نے ترجمہ کاری کے انداز کو سادہ لغت سے لے کر انتہائی پیچیدہ مصنوعی ذہانت پر مبنی نظام تک پہنچا دیا۔ ابتدا میں سادہ لغت اور قواعد پر مبنی نظامات تھے، جنہیں Rule-Based Machine Translation (RBMT) کہا جاتا ہے۔ ان نظاموں میں ہر لفظ اور اس کا متبادل لغت میں درج ہوتا، اور نحو کی بنیاد پر جملے کا ترجمہ کیا جاتا۔ یہ نظام لسانی اصولوں پر سختی سے کاربند رہتے، مگر ان کی محدودیت یہی تھی کہ وہ غیر متوقع یا نئی زبان کی تشکیل سے قاصر تھے۔ Kay کے کام سے متاثر ہو کر Constraint-Based Formalism نے لغوی قواعد کے ذریعے ترجمہ کی کوشش کی، مگر جیسے ہی معنی اور سیاق کی گہرائی درکار ہوئی، ان نظاموں کی خامیاں ظاہر ہونے لگیں۔

RBMT سے آگے بڑھ کر ماہرین نے Corpus-Based Machine Translation (CBMT) پر کام شروع کیا۔ اس میں متن کا ترجمہ کارپس یعنی تحریری مواد کے بڑے ذخیرے سے کیا جاتا ہے۔ یہ طریقہ دو سطحوں پر انجام پاتا ہے: ایک براہ راست، جس میں متن کا ترجمہ فوراً دستیاب کارپس سے کیا جاتا ہے؛ دوسرا بالواسطہ، جس میں لغت، شماریاتی ڈیٹا اور قواعد کی مدد سے اندازہ لگایا جاتا ہے کہ دیے گئے جملے کا متوقع ترجمہ کیا ہوگا۔ CBMT کے ذیلی ماڈلز میں Example-Based، Statistical اور Connectionist طریقے نمایاں ہوئے۔ Statistical Machine Translation (SMT) خاص طور پر Google Translate کی ابتدائی شکل میں اہم رہا، جہاں لاکھوں جملوں کی تربیت کے ذریعے الگورتھم متوقع ترجمے فراہم کرتا تھا۔

اسی تسلسل میں Knowledge-Based Machine Translation (KBMT) کا تصور سامنے آیا، جو مصنوعی ذہانت کا گہرا استعمال کرتا ہے۔ یہ نظام صرف ترجمہ نہیں کرتا بلکہ علمی فہم (semantic understanding) کے ذریعے سیاق، اصطلاحات اور جملوں کی ساخت کو معنیاتی پیرائے میں دیکھتا ہے۔ کارنیگی میلن یونیورسٹی کی ٹیم نے KBMT میں قابل ذکر کام کیا۔ ان کے مطابق مشینی ترجمہ کاری کو اس وقت تک مکمل نہیں سمجھا جا سکتا جب تک وہ زبان کی نحوی اور معنوی ساخت کے ساتھ ساتھ ثقافتی اور سیاقی پیرایوں کو بھی نہ سمجھے۔ یہی وجہ ہے کہ مشینی ترجمہ کار کو "ذہین نظام" قرار دیا جانے لگا۔

اس دوران کچھ ماڈلز ایسے بھی سامنے آئے جو ترجمہ کاری کو مختلف صارفین کی ضروریات کے مطابق ڈھالتے تھے۔ مثلاً MT for Watchers (MT-W) ان افراد کے لیے تھا جو صرف ترجمہ دیکھنا چاہتے تھے۔ یہ محض لغت پر انحصار کرتا اور ایک سادہ "لغوی ترجمہ" فراہم کرتا۔ MT for Revisors (MT-R) نظرثانی کرنے والوں کے لیے تھا جو مشینی ترجمے کو انسانی تدوین کے ذریعے قابل فہم بناتے۔ Logos System کی مثال یہاں دی جا سکتی ہے، جس نے پہلے مسودے کی تیاری کے بعد انسانی ترمیم کی راہ ہموار کی۔ MT for Translators (MT-T) وہ نظام تھا جو انسانی مترجم کو لغات، اصطلاحات اور تھیسارس کی مدد فراہم کرتا، جیسے کہ Trados یا Eurolang۔

بعد ازاں، MT for Authors (MT-A) جیسا ماڈل سامنے آیا، جو مصنف کو تحریر کے ساتھ ساتھ ترجمہ کرنے میں مدد دیتا ہے۔ فرانسیسی نظام TITUS اور جاپانی نظام N-Tran اسی تصور کے تحت بنائے گئے، جہاں مصنف اپنی زبان میں لکھتا جاتا اور سسٹم فوراً اس کا ترجمہ فراہم کرتا۔ اس طرز کی ترجمہ کاری ادارہ جاتی سطح پر بہت مؤثر ثابت ہوئی، خاص طور پر جب یکساں مواد کو ایک سے زیادہ زبانوں میں پیش کرنا مقصود ہو۔

ترجمہ کاری کے ان تمام ماڈلز میں **ترجمے کے تین بنیادی انداز** واضح طور پر موجود رہے: انتقالی انداز(Transfer Model)، بین اللسانی ماڈل (Interlingua Model)،اور لغوی انداز(Lexicalist Model) ۔انتقالی انداز میں ترجمہ کاری کو تین مراحل میں تقسیم کیا جاتا ہے: تجزیہ، منتقلی،اور ترکیب۔ہر مرحلہ خود ایک علیحدہ سافٹ ویئر ماڈیول کے تحت کام کرتا ہے۔اس طرز کی بہترین مثال 1970ء میں یورپی کمیشن کے لیے تیار کردہ Systran نظام ہے، جسے Peter Toma نے ڈیزائن کیا۔ Systran نظام آج بھی یورپی یونین کے کئی اداروں میں استعمال ہو رہا ہے۔

Interlingua Modelاس تصور پر مبنی تھا کہ ایک "درمیانی زبان" ہو جو ماخذ اور ہدف زبان دونوں سے آزاد ہو۔اس نظریے کی بنیاد ستر ہویں صدی کے اس خواب پر تھی کہ دنیا کے تمام انسان ایک مشترکہ زبان استعمال کریں۔یہی وجہ ہے کہ DLT System میں اسپرانتو کو Interlingua کے طور پر استعمال کیا گیا۔اگرچہ یہ تصور علمی طور پر بلند تھا، مگر عملی سطح پر مختلف زبانوں کے درمیان تنوع نے اس پر کئی سوالات اٹھا دیے۔

آج کی مشینی ترجمہ کاری ان تمام ماڈلز کا امتزاج ہے، جسے Hybrid Machine Translation (HMT) کہا جاتا ہے۔یہ نظام قواعد، شماریات، کارپس اور علمی فہم کو یکجا کر کے بہتر نتائج فراہم کرتا ہے۔گوگل ٹرانسلیٹ نے پہلے SMT اپنایا، پھر Neural Machine Translation (NMT) پر منتقل ہو گیا، جو ترجمہ کو پورے جملے، بلکہ پیراگراف کی سطح پر دیکھتا ہے۔

Wilksکے الفاظ میں:

> "The future of MT lies not in choosing between rules and statistics, but in combining them wisely with knowledge representation and neural inference."[5]

اسی وژن کے تحت جدید سسٹمز جیسے DeepL [6]، Microsoft Translator[7]اور Amazon Translate [8]بنائے گئے، جو ترجمہ کاری کو نہ صرف درست اور تیز کرتے ہیں بلکہ اسے زبان کی "فہم" کے قریب تر لاتے ہیں۔

عصرِ جدید میں مشینی ترجمہ کاری نے **نیورل نیٹ ورک اور مصنوعی ذہانت** کے انضمام سے تیزی اختیار کی۔ **DeepL**ایک مشہور Neural Machine Translation (NMT)سروس ہے، جو شروع میں سات یورپی زبانوں کے لیے متعارف ہوئی مگر اب اس کی پہنچ 35 زبانوں تک ہو چکی ہے۔اس ماڈل کی بنیاد Transformerمعماری پر ہے اور اسے Linguee ڈیٹابیس سے تربیت دی گئی ہے، جس نے ترجمے کو نہایت قدرتی اور معنی خیز بنایا ہے۔اس کا ترجمہ اکثر انسان جیسا محسوس ہوتا ہے، خاص کر یورپی زبانوں کے لیے۔

اسی طرح، Microsoft Translatorنے بھی NMT اپنایا ہوا ہے اور اسے Cognitive Servicesکے تحت متعدد پلیٹ فارمز—جیسے Bing،Office،Skype Translator،وغیرہ—میں شامل کیا گیا ہے۔اس کی طاقت enterprise سطح پر اپنی مرضی کے ماڈلز بنانے اور real-timeبات چیت کے دوران ترجمہ فراہم کرنے میں ہے۔

Amazon Translateایک AWS کا NMT سروس ہے جس نے درجنوں زبانوں کے لیے real-time اور large-scale batch ترجمہ فراہم کرنے کی صلاحیت فراہم کی ہے۔ یہ مخصوص صنعتوں یا ڈومینز کے مطابق ماڈل کو customize کرنے کا عملی حل بھی پیش کرتا ہے Timekettlelingohub.com۔

DeepLاپنی درستی اور معنوی تاثیر کے لیے خاص طور پر سراہا جاتا ہے۔یہ مخصوص industry-specific glossaries کو integrate کرنے، document layoutکو برقرار رکھنے، اور ترجمہ کی ٹون کو برقرار رکھنے کی قابلیت رکھتا ہے — جو کہ ترجمہ کو زیادہ انسانی محسوس کرواتا ہے TechRadarsmartling.comPhrase۔

اگرچہ یہ بین الاقوامی نظام زبردست ہیں، کم وسائل والی زبانوں (low-resource languages) جیسے اردو میں مشینی ترجمہ ابھی بھی چیلنج ہے۔جلدی میں کوئی پیشرفت نظر نہیں آ رہی، کیونکہ اردو کے لیے ترجمہ کارپس اور تربیتی ڈیٹا محدود ہے، جس کی بنا پر NMTماڈلز کی کارکردگی دوسری زبانوں کے مقابلے میں کمزور رہتی ہے ACL Anthologyjournals.uol.edu.pk۔

تحقیقی مقالوں میں (FConv-NN) Fully Convolutional Neural Network پر مبنی Urdu-English ماڈل کو استعمال کرتے ہوئے کچھ بہتری کی طرف اشارہ کیا گیا ہے، جو جملوں میں long-range dependencies کو بہتر طریقے سے سمجھتا ہے ResearchGate۔ اسی طرح، LSTM Encoder-Decoder بہودی میں پیشرفت کا ثبوت ہے، لیکن اس کے باوجود اردو کے لیے ترجمہ کی موازنہ قابلیت محدود ہے

اردو-انگریزی ترجمہ میں (FConv-NN) Fully Convolutional Neural Network کا تجربہ کیا گیا [14]۔ اسی طرح LSTM-Decoder ماڈلز کی مدد سے اردو ترجمہ کو بہتر بنانے کی کوشش کی گئی [13]۔

ایک اور تحقیقی اقدام Example-Based MT کے حوالے سے پاکستان میں سامنے آیا، جہاں interactive English-to-Urdu MT سسٹم بنایا گیا، جس میں انسانی مداخلت سے ترجمہ کو زیادہ معیاری بنایا گیا ResearchGate+1۔

Pakistan میں Example-Based Approach پر مبنی interactive English-to-Urdu MT سسٹم بھی تیار کیا گیا [15]۔

اسی تسلسل میں، خلائی صنعت اور یونیورسٹیوں نے بھی رول ادا کیا؛ لیکن عالمی معیار پر اردو کی MT تحقیق ابھی ابتدائی مراحل میں ہے، جس میں بہتری کی گنجائش اور وسائل کی ضرورت واضح ہے۔

مصنوعی ذہانت اور Neural MT نے عالمی سطح پر مشینی ترجمہ کاری کو نئے افق دے دیے ہیں، جس میں **DeepL**، **Microsoft Translator** اور **Amazon Translate** جیسے سروسز نمایاں ہیں۔

ACL Anthology اور JCSA جیسے تحقیقی جرائد میں اردو MT پر تحقیقی مضامین کی اشاعت جاری ہے [16]۔

مگر اردو جیسی زبانوں کے لیے MT کی کامیابی کم ہی ہے، کیونکہ تربیتی ڈیٹا کا فقدان گہرائی میں پیش آتا ہے۔ چند تحقیقی ماڈلز (جیسے FConv-NN، LSTM-Decoder، Example-Based Systems) امید کی کرن رکھتے ہیں، مگر مستقل اور موثر نظام اب بھی دور ہے۔

حوالہ جات:

1. Yorick Wilks, *Machine Translation: Its Scope and Limits* (London: Springer, 2009), 3, 158.
2. John Wilkins, *An Essay Towards a Real Character and a Philosophical Language* (London: Royal Society, 1668).
3. W. John Hutchins and Harold L. Somers, *An Introduction to Machine Translation* (London: Academic Press, 1992).
4. ALPAC (Automatic Language Processing Advisory Committee), *Languages and Machines: Computers in Translation and Linguistics* (Washington, DC: National Academy of Sciences, 1966).
5. Yorick Wilks, *Machine Translation: Its Scope and Limits* (London: Springer, 2009), 3, 158.
6. "DeepL Translator." *Wikipedia*. Last modified August 17, 2025. https://en.wikipedia.org/wiki/DeepL_Translator.
7. "Microsoft Translator." *Wikipedia*. Last modified August 17, 2025. https://en.wikipedia.org/wiki/Microsoft_Translator.
8. "5 Best Machine Translation Software for Business." *Timekettle*. Accessed August 18, 2025. https://www.timekettle.co/en-ca/blogs/tips-and-tricks/5-machine-translation-software.

9. "Neural Machine Translation and the Future of Multilingual AI." *TechRadar*. Accessed August 18, 2025. https://www.techradar.com/news/neural-machine-translation-future.

10. "Best AI Translation Tools in 2025: Smartling, Phrase, and Beyond." *Smartling.com*. Accessed August 18, 2025. https://www.smartling.com/resources/ai-translation-tools-2025.

11. "Industry-Specific MT Customization." *LingoHub*. Accessed August 18, 2025. https://lingohub.com/blog/industry-specific-mt.

12. "Document Layout and Glossary Integration in DeepL." *Phrase.com*. Accessed August 18, 2025. https://phrase.com/blog/deepl-document-layout.

13. Ubaid Ullah et al., "Towards Urdu English Neural Machine Translation: A Comparative Study of Transformer and LSTM Architectures." *Journal of Computer Science Applications* 12, no. 4 (2024): 95–106. .https://journals.uol.edu.pk/JCSA/article/download/3536/1733/18524

14. "Improved Urdu English Neural Machine Translation with a Fully Convolutional Neural Network Encoder." *ResearchGate*. Accessed August 18, 2025. https://www.researchgate.net/publication/384455318.

15. "Interactive English to Urdu Machine Translation Using Example Based Approach." *ResearchGate*. Accessed August 18, 2025. https://www.researchgate.net/publication/41392016.

16. "Low Resource Neural Machine Translation Workshop." *ACL Anthology*. Accessed August 18, 2025. https://aclanthology.org/events/loresmt.