مشینی ترجمہ: تاریخ، حال اور مستقبل
اردو کے تناظر میں ایک مطالعہ

ڈاکٹر حافظ صفوان محمد چوہان/ڈاکٹر ظہیر احمد/انجینئر سارہ سلیم

This article discusses the history of machine-aided & computer-aided translation service, its current scenario with categoric reference of available online translation services, and its future. Where do the Urdu and its script stand in this situation are taken into account on reality and pure technical basis.

Environment of this article is made reasonably digestable for the people of Urdu literature so as to retain their attention in these pages of stark technical discussion.

اصولی اور تکنیکی طور پر اِس مقالے کا دائرۂ کار کسی بھی زبان سے اردو اور اردو سے کسی بھی زبان میں کسی بھی قسم کا متن ترجمہ کرنے پر پھیلا ہوا ہے۔ تاہم مثالوں کی حد تک یہ صرف اردو-انگریزی اور انگریزی-اردو سے بحث کرتا ہے۔

یہ مقالہ فروری ۲۰۱۰ء کے آخری ہفتے میں مکمل ہوا۔ اِس میں شامل معلوماتِ عامہ کو صرف اِنہی تاریخوں تک درست سمجھا جائے۔

کلیدی الفاظ

مشین ریڈیبل اردو، کارپس، ماخذ جملہ (Source Sentence)، انسانی زبان (Natural Language)، ماخذ زبان (Source Language)، مطلوبہ زبان (Target Language)، مترجم (Human Translator)، شعبے کی زبان (Term-base)، لفظیات (Lexis)، متن (Discourse)، نقلِ حرفی (Transliteration)، کمپیوٹیشنل گرامر (Computational Grammar)، فیصلہ کرنے کی صلاحیت (Decision-making)، مصنوعی ذہانت (Artificial Intelligence)۔

☆اِس مقالے میں لفظ زبان زیادہ تر انسانی زبان کے مفہوم میں استعمال کیا گیا ہے۔

☆ترجمہ یا زبان کا ترجمہ سے بھی عموماً مراد زبانوں کے کسی جوڑے (مثلاً اردو، انگریزی) کا باہم ترجمہ ہے۔

مخففات

ٹرانسلیشن میموری:	TM: Translation Memory
مشین ٹرانسلیشن، مشینی ترجمہ:	MT: Machine Translation
ٹریڈوس:	TRADOS: Translator for DOS
مصنوعی ذہانت:	AI: Artificial Intelligence
شماریاتی مشینی ترجمہ:	SMT: Statistical Machine Translation
ناموجود الفاظ:	OOV: Out of Vocabulary Words
رک:	رجوع کیجیے
مرکز تحقیقات اردو پاکستان:	CRULP: Centre of Research in Urdu Language Processing
مقتدرہ قومی زبان پاکستان:	NLA: National Language Authority, Islamabad, Pakistan

اصطلاحات

شعبے کی زبان:	Term-base
عالمی ترجمہ کار:	Universal Translator
کمپیوٹیشنل لسانیات:	Computational Linguistics
اطلاعیات:	Informatics
بالائے لغت معنی:	Ultra-dictionary meaning
پلیٹ فارم سے نا وابستہ:	Platform-Independent
روزمرہ زبان/بول چال:	Contemporary [use of] Language
انسانی/فطری زبان:	Natural Language
کارپس/امثال گھر/قاموس الامثال:	Corpus
لفظیاتی تجزیہ:	Lexical Analysis
ڈیٹا کا ذخیرہ/مال خانہ:	Repository
موقع:	(Internet) Site

☆**تجارتی نشانات:** اس مقالے میں Microsoft، Unicode، Google، Yahoo، Facebook، وغیرہ الفاظ بار بار استعمال کیے گئے ہیں؛ یہ الفاظ مختلف اداروں کے تجارتی نشانات (ٹریڈ مارک) ہیں۔

0: تعارف

مشینی ترجمے سے مراد ایک انسانی زبان (ماخذ زبان) کے متن کا دوسری انسانی زبان (مطلوبہ زبان) میں سافٹ ویئر کی مدد سے

ترجمہ کرنا ہے۔ مشینی ترجمہ شعبۂ کمپیوٹیشنل لسانیات کی ایک شاخ ہے۔

مشینی ترجمے کا سادہ ترین انداز یہ ہے کہ ایک زبان کے الفاظ کو دوسری زبان کے الفاظ سے صرف بدل دیا جائے (Substitution)۔ کارپس کی تکنیک کے استعمال سے نسبتۂ پیچیدہ ترجمے جن میں زبان کی ساخت، جملوں کی بنت اور پہچان، اور محاوروں اور دیگر اجزائے کلام وغیرہ کو ملحوظ رکھا گیا ہو، بہتر انداز میں کیے جا سکتے ہیں۔

1: مشینی ترجمے کی تاریخ

زبانوں کے آڑے کھوڑ (Barriers) کو میکانکی آلات کے ذریعے سے توڑنے کا خیال سترھویں صدی میں پیدا ہوا جب لاطینی زبان مردہ ہو چکی تھی اور سائنسی تحقیقات کو ایسے الفاظ میں دوسری زبانوں میں پیش کرنے کی ضرورت محسوس ہونے لگی تھی جنھیں دوسری زبانیں بولنے والے لوگ سمجھ سکیں۔ ایک ''عالمی زبان'' کے خیال کی تہہ میں یہ دونوں باتیں شامل تھیں کہ پوری دنیا کے لوگ ایک دوسرے سے گفتگو کر سکیں اور سائنسی معلومات بھی درست طریقے سے دوسری زبانوں والوں کو پہنچائی جا سکیں۔

ہر مشینی ترجمہ کمپیوٹر سے یا کمپیوٹر کی مدد سے ہو، یہ ہرگز ضروری نہیں۔ چنانچہ مشینی ترجمے کی تاریخ کو دو بڑے حصوں میں تقسیم کیا جا سکتا ہے: کمپیوٹر کے دور سے پہلے، اور کمپیوٹر کی آمد کے بعد۔

1.1: مشینی ترجمہ—کمپیوٹر کی آمد سے پہلے

۱۶۲۹ء میں فرانسیسی فلاسفر اور ماہرِ طبعیات René Descartes (1596–1650) نے ایک عالمی زبان کا تصور پیش کیا جس میں معلوم زبانوں کے ملتے جلتے خیالات والے الفاظ کو ایک ہی علامت (Code) سے ظاہر کرنے کی تجویز تھی۔ مختلف زبانوں کے مماثل الفاظ کو ایک ہی علامت دینے کا ایسا ہی نظریہ جرمن فلاسفر اور علامۂ دہر Gottfried Wilhelm Leibniz (1646–1716) نے بھی پیش کیا۔ ۱۶۶۸ء میں John Wilkins نے اِس بارے میں بہتر نظریہ پیش کیا کہ مختلف زبانوں کے مترادف الفاظ کو ایک ہی کوڈ سے ظاہر کیا جائے۔ تاہم اِن سب نظریات میں مستقل میکانکی مشین بنانے کا کوئی خیال شامل نہ تھا بلکہ انسانوں ہی سے یہ میکانکی عمل کرانے کی بات تھی۔ یہ نظریات محض تصورات ہی رہے اور انھیں صورت ملنا بیسویں صدی ہی میں شروع ہو سکا۔

''ترجمہ مشین'' کو پیٹنٹ کرانے کی پہلی درخواست ۱۹۳۳ء میں جمع کرائی گئی۔ ۲۲/ جولائی ۱۹۳۳ء کو ایک فرانسیسی انجینئر Georges Artsrouni کو ''میکانکی دماغ'' بنانے کا پیٹنٹ دیا گیا۔ یہ ایک سادہ سا خودکار ذولسانی لغت تھا جو کاغذی ٹیپ (Paper-tape) کے ذریعے کام کرتا تھا۔ ایک اور نظریہ جو ۱۹۳۷ء میں ایک روسی استاد اور عالم Petr Petrovich Troyanskii (1894–1950) نے پیش کیا، نسبتۂ زیادہ تفصیلی تھا۔ اِس میں ایسپرانتو زبان (Esperanto) کی بنیاد پر تیار کردہ ایک ذولسانی لغت کے ساتھ ساتھ گرامر کے بین اللسانی قواعد سے نمٹنے کا طریق کار بھی تھا۔ ترجمے کے اِس نظام کو تین سطحوں میں بانٹا گیا تھا: پہلے حصے میں صرف ماخذ زبان جاننے والا ایک آدمی ماخذ متن کے الفاظ کو اُن کی سادہ حالتوں (Base-forms) اور نحوی ترتیب میں لکھ کر منطقی تجزیے کی علامات (Logical Analysis Symbols) میں تبدیل کر دیتا تھا؛ دوسری سطح میں مشین اِن علامات کو مطلوبہ زبان کی ایسی ہی علامات میں ترجمہ کر دیتی تھی؛ اور تیسری سطح پر صرف مطلوبہ زبان جاننے والا ایک آدمی اِن علامات یعنی آؤٹ پُٹ کو اپنی زبان کے مطابق ٹھیک ٹھاک کر کے

پیش کر دیا تھا۔ ٹرائسکی کی یہ سکیم جو بنیادی طور پر اگرچہ صرف لغت ہی کو خودکار بنانے کی کوشش تھی اور جس میں دو انسان (مترجم) مستقلاً چاہیے ہوتے تھے، ۱۹۵۰ء کے اواخر تک پردۂ گمنامی میں رہی یہاں تک کہ کمپیوٹر عام ہو گئے۔ ٹرائسکی کو بجا طور پر بابائے مشین ٹرانسلیشن (Father of Machine Translation) تسلیم کیا جاتا ہے۔

مشینی ترجمے کا استعمال دوسری جنگِ عظیم میں سامنے آیا جب ہٹلر نے اینگما (Wehrmacht Enigma) استعمال کرتے ہوئے کوڈز کے ذریعے سے پیغامات کی ترسیل کا کام کر لیا۔ ویسے اینگما مشینیں ۱۹۲۰ء سے استعمال میں تھیں۔

1.2: مشینی ترجمہ — کمپیوٹر کی آمد کے بعد (پہلا دور: ۱۹۴۶ء سے ۱۹۶۰ء تک)

الیکٹرانک کیلکولیٹر کی ایجاد کے فوراً بعد کمپیوٹر کے ذریعے انسانی زبانوں کے ترجمے پر تحقیقات کی باتیں شروع ہوئیں۔ ۱۹۴۶ء میں Andrew Donald Booth اور کچھ اور لوگوں نے ڈیجیٹل کمپیوٹر کے ذریعے انسانی زبانوں کے ترجمے کا خیال سب سے پہلے پیش کیا۔ مارچ ۱۹۴۷ء میں امریکی محقق Warren Weaver نے اِس سلسلے میں ایک خط لکھا۔ پھر اُس نے کچھ مقالات لکھے اور کمپیوٹر کے ذریعے مشینی ترجمے کی باقاعدہ تجویز پہلی بار جولائی ۱۹۴۹ء میں پیش کی۔ اُس کی تجاویز کی بنیاد انفارمیشن تھیوری، دوسری جنگِ عظیم میں اینگما مشین کے کوڈ توڑنے (Code-breaking) میں کامیابیاں پانے، اور اِس قیاس پر بنی تھی کہ سب انسانی زبانیں چند عالمگیر اصولوں پر کام کرتی ہیں (possibilities of language universals)۔

مشینی ترجمے کی پہلی کانفرنس Yehoshua Bar-Hillel کی کوششوں سے ۱۷ تا ۲۰/ جون ۱۹۵۲ء کو امریکہ کے Massachusetts Institute of Technology میں Rockefeller Foundation کے تعاون سے منعقد ہوئی۔ اِس کانفرنس میں مشینی ترجمے پر تحقیقات میں دلچسپی لینے والے اٹھارہ سائنس دانوں نے شرکت کی۔ یہاں پڑھے گئے مقالات میں مشینی ترجمے کے مسائل پر کئی پہلوؤں سے روشنی ڈالی گئی اور بہت سی تجاویز پیش کی گئیں۔

اِن تجاویز کے ساتھ ہی امریکہ کی کئی یونیورسٹیوں میں زبانوں (اور اُن کے ترجمے) پر تحقیق کا کام درحقیقت شروع ہوا۔ چنانچہ Georgetown-IBM تجربات کے نتیجے میں ۷/ جنوری ۱۹۵۴ء کو نیویارک میں IBM کے صدر دفتر میں کامل طور پر خودکار مشینی ترجمے کی پہلی عوامی نمائش ہوئی جس کی اخبارات اور میڈیا میں خوب تشہیر ہوئی اور جس میں عوام نے بھی بہت دلچسپی ظاہر کی۔ اُس وقت پیش کی گئی ترجمہ مشین کو آج تو صرف کھلونا ہی کہا جائے گا کیونکہ اِس میں صرف ۲۵۰ الفاظ تھے اور گرامر کے چھے بنیادی اصول، جن کے ذریعے انتہائی توجہ سے منتخب کیے گئے روسی زبان کے ۴۹ جملوں کو انگریزی میں ترجمہ کیا گیا تھا۔ یہ جملے زیادہ تر شعبۂ کیمیا سے متعلق تھے۔ بہرحال اِس کامیابی سے یہ ضرور ہوا کہ صرف امریکہ ہی نہیں بلکہ پوری دنیا میں مشینی ترجمے کے کام میں پھریری آ گئی اور اِن تحقیقی منصوبوں کے لیے بھاری رقوم مختص کی جانے لگیں۔ اِس ترجمے کے کرنے والے سائنسدان بہت پُر امید تھے کہ آئندہ تین سے پانچ سال کے اندر مشینی ترجمے کا مسئلہ ایک حل شدہ سوال ہوگا۔

حقیقی ترقی کی رفتار بہرحال بے حد سست تھی۔ اِن ابتدائی کوششوں میں مطلوبہ زبان میں ترجمے کے حتمی آؤٹ پُٹ کے لیے بڑی بڑی دولسانی لغات اور الفاظ کی ترتیب کو درست کرنے کے لیے بے شمار دستی قواعد استعمال کیے جاتے تھے۔ رفتہ رفتہ معلوم ہوا کہ یہ طریقے چلنے

کے نہیں۔ چنانچہ اِن ترجموں کا معیار بہتر بنانے کے لیے ساختیاتی لسانیات (Generative Linguistics) اور تحویلی گرامر (Transformational Grammar) وغیرہ جیسی کئی طرح کی لسانیاتی کوششیں کی گئیں۔

اِس دوران میں جگہ جگہ آپریشنل سسٹم بھی لگا لیے گئے۔ امریکی فضائیہ نے IBM کا تیار کردہ ایک سسٹم اور امریکی اٹامک انرجی کمیشن نے جارج ٹاؤن یونیورسٹی کا تیار کردہ ایک سسٹم اپنا لیا۔ اِن کا مقصد زیادہ تر، یہی تھا کہ روسی زبان میں کی جانے والی گفتگو کو انگریزی میں سمجھا جا سکے۔ لیکن اِن ترجمہ کار مشینوں کی کارکردگی بہت اچھی نہ تھی اور یہ اپنے استعمال کنندگان کی بہت سی ضرورتیں پوری نہ کر سکتی تھیں۔ خصوصاً اِن کی سپیڈ بے حد کم تھی۔

۱۹۵۰ء کی دہائی کے اواخر میں امریکی حکومت نے Yehoshua Bar-Hillel کو مشینوں کے استعمال سے ایک ''کامل خودکار اعلیٰ درجے کا ترجمہ کار'' (Fully Automatic High Quality Translator جس کا مخفف FAHQT ہے) بنانے کے امکانات کا جائزہ لینے کا کام سونپا۔ اُس نے جملوں میں پائے جانے والے معنیاتی ابہام (Semantic Ambiguity) یعنی ذو معنی الفاظ والے جملوں کا درست ترجمہ کرنے کے میدان پر کام کیا۔ اُس کے پیشِ نظر ایسے جملے رہے جن میں ایسے الفاظ ہوں جن کا مطلب انسان تو سمجھ سکے لیکن جن کا ترجمہ کرنے میں مشین چکرا جائے، مثلاً pen کا معنی قلم ہو یا کٹہرا، یا کوئی اور۔ اُس کا دعویٰ تھا کہ جب تک کوئی ''عالمگیر انسائیکلوپیڈیا'' نہ بن جائے، مشین اِس قسم کے مسائل سے نہیں نپٹ سکتی۔ یہ الگ بات ہے کہ آج اِس قسم کے ابہام سے نبرد آزما ہونے کے لیے مشینی ترجمے کے سافٹ ویئر میں صرف چند لائنوں کا ایک کوڈ لکھنا پڑتا ہے۔

1.2.1: مشینی ترجمہ — دوسرا دور: ۱۹۶۱ء سے ۱۹۸۰ء تک

۱۹۶۰ء کی دہائی میں روس اور امریکہ دونوں میں مشینی ترجمے کا کام انگریزی - روسی ترجمے کے گرد گھومتا رہا۔ ترجمے کے متن زیادہ تر سائنسی اور تکنیکی مقالے ہوتے تھے۔ ڈھیلے ڈھالے ترجمے سے یہ اندازہ کر لیا جاتا تھا کہ مقالے میں کیا لکھا ہے۔ اگر کسی مقالے کے بارے میں محسوس ہوتا کہ یہ سیکورٹی کے مسائل کی دلچسپی کا ہے تو اُسے کسی ماہر مترجم کے پاس مکمل ترجمے کے لیے بھجوا دیا جاتا۔

نومبر ۱۹۶۶ء میں مشینی ترجمے کی تحقیقات پر ایک شدید ضرب لگی جب John R Pierce کی سربراہی میں تشکیل دی گئی سات سائنسدانوں پر مشتمل ایک کمیٹی ALPAC (Automatic Language Processing Advisory Committee) نے اپنی تحقیقاتی رپورٹ شائع کی۔ یہ کمیٹی امریکی حکومت نے ۱۹۶۴ء میں بنائی تھی۔ امریکی حکومت اور مشینی ترجمے پر تحقیقات کی مالی سرپرستی کرنے والے سرمایہ دار اِس بات پر فکرمند تھے کہ بہت خرچ ہو جانے کے باوجود بھی کام آگے نہیں بڑھ رہا۔ رپورٹ میں بتایا گیا تھا کہ مشینی ترجمہ مہنگا زیادہ ہے، درست کم ہے اور انسانی مترجم کی نسبت کم رفتار ہے۔ یہ بھی بتایا گیا کہ خرچ خواہ کتنا بھی کر لیا جائے، لگتا نہیں کہ مستقبل قریب میں مشین سے کیے گئے ترجمے کا معیار انسانی مترجم تک پہنچ پائے۔ رپورٹ میں یہ سفارش البتہ ضرور کی گئی تھی کہ ترجمے کے لیے مددگار سافٹ ویئر — جیسے خودکار لغات — اور کمپیوٹیشنل لسانیات وغیرہ پر تحقیق جاری رکھی جائے۔ اِس رپورٹ پر شدید اعتراضات ہوئے اور جانبداری کے الزامات بھی لگے۔

ALPAC کی اِس بدنام زمانہ رپورٹ کی اشاعت سے مشینی ترجمے کی تحقیقات نے پٹخنی کھائی۔ امریکہ میں تو یہ تحقیقات تقریباً

ایک عشرے کے لیے بالکل ہی رک گئیں۔ البتہ روس اور برطانیہ کے علاوہ کناڈا، فرانس اور جرمنی میں یہ کام چلتا رہا۔ اِن ملکوں کی تحقیقات کے نتیجے میں ۷۶-۱۹۷۵ء میں مانٹریال یونیورسٹی میں بننے والا ایک پروگرام METEO System کناڈا میں انگریزی-فرانسیسی اور فرانسیسی-انگریزی مشینی ترجمے کے ساتھ ۱۹۷۷ء سے ستمبر ۲۰۰۱ء تک موسمیاتی رپورٹ دیتا رہا ہے۔

۱۹۶۰ء کی دہائی میں مشینی ترجمے کی تشکیل کے سارے ہی کام جنگی (یا دفاعی) ضرورتوں کے تناظر میں کیے گئے جب کہ ۱۹۷۰ء کی دہائی میں تکنیکی اور کاروباری دستاویزات کے کم قیمت تراجم پر زور رہا۔ پھر گلوبلائزیشن میں اضافے کی وجہ سے کناڈا، یورپ اور جاپان میں مشینی ترجمے کی طلب میں اضافہ ہوا۔

1.2.2: مشینی ترجمہ — تیسرا دور: ۱۹۸۱ء سے ۱۹۹۰ء تک

۱۹۸۰ء تک مشینی ترجمے کے لیے لگائی گئی مشینوں کی تعداد میں بہت اضافہ ہوا۔ مزید چھوٹے کمپیوٹروں کے آجانے سے چھوٹے کاروباریوں کی کم قیمت میں ترجمہ کرانے کی مانگ بڑھی۔ اِس دہائی میں جاپان میں مشینی ترجمے پر بڑا کام ہوا۔ انگریزی سے اور انگریزی میں مشینی ترجمہ کے لیے جاپان کی چھوٹی بڑی سبھی کمپنیاں میدان میں کود پڑیں۔

اِس دہائی میں مشینی ترجمے پر تحقیقات زیادہ تر لیکن 'بالواسطہ' درمیانی لسانیاتی صورتوں پر انحصار کرتی رہیں جس میں الفاظ کی تصریفی شکلوں (Morphology) اور ترکیب نحوی (Syntax) اور الفاظ کی معنویت (Semantics) میں بیک وقت تطابق ہو۔

۱۹۸۰ء کے اواخر تک مشینی ترجمے کے کئی جدید طریقے سامنے آگئے۔ ایک طریقہ IBM کا تھا جس کی بنیاد الفاظ کے شماریاتی ماڈل (Statistical Method) پر تھی۔ ایک اور طریقہ جسے آج مثالی ترجمہ کاری (Example-based Machine Translation) کہتے ہیں، ترجمہ شدہ مثالی جملوں پر انحصار کرتا تھا۔ اِن دونوں طریقوں کا ایک خلا جملوں کی نحوی اور الفاظ کی معنوی تراکیب کے لیے قواعد کا نہ ہونا اور اِن کی بجائے بڑے بڑے متون پر مشتمل کارپس پر اندھا دھند اعتبار کرنا تھا۔

1.2.3: مشینی ترجمہ — چوتھا دور: ۱۹۹۱ء تا حال

۱۹۹۰ء کی دہائی میں مشینوں کے ذریعے آوازوں کی پہچان اور آواز پر تحقیقات میں بڑی کامیابیاں حاصل ہوئیں اور آواز سے ترجمے (Speech to Text) پر تحقیق شروع ہوئی۔ کم قیمت اور تیز رفتار کمپیوٹروں کے آجانے سے مشینی ترجمے کے استعمال میں بے حد اضافہ ہوگیا۔ مشینی ترجمے کا کام Mainframe کمپیوٹروں سے ہٹ کر چھوٹے ذاتی کمپیوٹروں (Personal Computers) پر آگیا۔ سب کمپنیوں نے اپنے سافٹ ویئر اِن کمپیوٹروں کے لیے ڈھالنا شروع کردیے۔ مشینی ترجمے پر نری تجرباتی تحقیق یعنی تحقیق برائے تحقیق کی بجائے ایسے سسٹم بنانے پر کام شروع ہوا جن کی عملی اور کاروباری افادیت ہو۔ ترجمے کی سہولت دینے والی سائٹیں اِنٹرنیٹ پر بھی ملنے لگیں۔

عام انسانی ضرورتوں کے لیے کمپیوٹر کی مدد سے ترجمے کا کام ۱۹۹۰ء کے بعد یعنی چھوٹے کمپیوٹروں کے عام ہونے پر شروع ہوا۔

2. مشینی ترجمے پر اردو کے لحاظ سے ایک نظر

یاد رکھنے کی ایک اہم بات یہ ہے کہ زبان قواعد کی پابند نہیں ہوتی اور نہ ہی کسی انسانی زبان کو مشین کا پابند کیا جاسکتا ہے۔ زبان بہت پہلے سے بنی ہوتی ہے اور اُس کے قواعد بعد ازاں گھڑے (یا اخذ کیے) جاتے ہیں۔ مثال لیجیے کہ اردو کے الفاظ میں جمع عام طور سے دو طرح کی

ہوتی ہے جیسے لفظ "کتاب" کی جمع ہے: "کتابیں" اور "کتابوں"۔ مشینی ترجمہ کار کو یہ سمجھانا اور بجھانا کہ جملے کی جنس کے اعتبار سے "کتابیں" یا "کتابوں" کہاں کہاں اور کیسے کیسے آئے گا، ایک طویل کمپیوٹر پروگرام لکھا جانے کا محتاج ہے اور جس سے مستثنیات کی فہرستیں منسلک کردیے کے باوجود بھی غلطی کا احتمال رہے گا۔ اور چونکہ لفظ "کتاب" عربی الاصل ہے جس کی اصولی جمع "کتب" بھی اردو میں عام رائج ہے (اور کلاسیکی ادبی متون میں جس کی جمع "کتاباں" بھی ملتی ہے)، لہٰذا یہ اندازہ لگانے میں عام سمجھ بوجھ کے آدمی کو بھی مشکل نہیں ہوگی کہ اردو ترجمہ کار سافٹ ویئر میں صرف واحد سے جمع (یا جمع سے واحد) بنانے ہی میں کیسی کیسی نزاکتیں، طوالتیں اور گھمن گھیریاں ہیں۔ اس کے برخلاف واحد/جمع کے لیے انگریزی میں چند مستثنیات کی ایک نسبۃً بہت ہی مختصر فہرست کے ساتھ صرف ایک ہی صورت ہوتی ہے اور جس میں جنس کا مسئلہ بھی نہیں ہوتا۔ عربی میں جیسے ایک مزید صورت، مثنیہ، بھی ہوتی ہے۔ مزید یہ کہ اردو میں جمع الجمع بھی عام پائی جاتی ہے جو دوسری زبانوں میں بہت کم ہے مثلاً "خبر" کی جمع "اخبار" اور "اخبار" کی جمع "اخبارات"۔ خیال رہے کہ یہ اردو کی خوبی ہے نہ کہ خرابی۔

مندرجہ بالا مثال میں اردو کے صرف ایک قاعدے یعنی جمع / واحد کا ذکر ہے اور وہ بھی محض اجمالی۔ اردو افعال کے صیغوں کا بھی یہاں کوئی ذکر مذکور نہیں جن کی تعداد ۲۵ ہے اور جملے کی نحوی اقسام (اسمیہ، فعلیہ، انشائیہ، خبریہ وغیرہ) کا بھی۔ ہر زبان کی طرح اردو میں بھی قواعد کی ایک لمبی فہرست ہے اور ان سب کو مشینی استعمال کے لیے کارگر بنانا پیچ در پیچ کام ہے۔ یہ کام مختلف جگہوں پر ہو رہا ہے۔ مرکز تحقیقات اردو پاکستان (CRULP) کی ویب سائٹ پر ایسا بہت سا مواد عام استفادے کے لیے اور بلا معاوضہ رکھا ہے جو مشینی ترجمے کی مختلف جہات سے تعلق رکھتا ہے اور جس سے ساری دنیا کے محققین زبان و لسانیات فائدہ اٹھا رہے ہیں۔ پیش آمدہ ضروریات کے تناظر میں تازہ بتازہ چیزیں اس سائٹ پر رکھی جاتی رہتی ہیں، اور اس کے حجم ہی نہیں بلکہ افادیت میں بھی روز بروز اضافہ ہو رہا ہے۔ مرکز تحقیقات اردو کی سائٹ پر موجود سہولیات کی ایک فہرست ملاحظہ کیجیے۔ مقتدرہ قومی زبان پاکستان (NLA) نے بھی جولائی ۲۰۰۹ء میں ایک پریس کانفرنس میں بلند بانگ الفاظ میں انگریزی - اردو مشینی ترجمے میں کامیابی حاصل کرنے اور اردو سے سائبر دور کا فاصلہ پاٹنے کا اعلان کیا ہے۔ ملاحظہ کیجیے:

www.thenews.com.pk/print1.asp?id=189086

یہ بات واضح ہو گئی کہ ایک زبان سے دوسری زبان میں کیا ہوا اچھا ترجمہ جو علمی اور دیگر ضروریات کے عین مطابق موافق ہو، کسی بھی قسم کے ترجمہ کار سافٹ ویئر کے ذریعے سے ایک انگلی کے اشارے سے حاصل نہیں ہو سکتا۔ ہاں، موجود سہولیات کو برتتے ہوئے بہتر سے بہتر کی طرف جایا جا سکتا ہے اور اردو سمیت دنیا کی بیشتر زبانوں کا یہ سفر جاری بھی ہے۔

اردو مشینی ترجمے پر کام شعبۂ انفارمیشن ٹیکنالوجی والوں کے لیے بہت بڑا چیلنج ہے۔ اس کی وجہ یہ بھی ہے کہ اردو دنیا کی سب زبانوں سے زیادہ رنگارنگی کی حامل ہے الفاظ کے اعتبار سے بھی اور قواعد کے اعتبار سے بھی۔ شاید اردو ہی وہ آخری زبان ہوگی جس کا اچھے نتائج کے ساتھ مکمل مشینی ترجمہ ہو سکے گا۔

2.1 ٹرانسلیشن سافٹ ویئر ذریعے اردو ترجمہ: یونی کوڈ متن کی ضرورت

انفارمیشن ٹیکنالوجی اور اطلاعیات (Informatics) کے سبھی کاموں میں ایک بات کو بغیر کسی بحث کے تسلیم کیا جاتا ہے کہ مشینی سہولیات سے فائدہ اٹھانے کے لیے ہر قسم کے متن کو مشین ریڈ ایبل حالت میں ہونا چاہیے۔ چنانچہ کمپیوٹر کے ذریعے ایک زبان سے دوسری

زبان میں ترجمہ کرنے کی پہلی ضرورت یہی ہے کہ متن مشین ریڈ ایبل حالت میں مہیا ہو؛ اِس ضرورت سے اردو کو بھی استثنا نہیں۔ لہٰذا اردو کے روایتی رسم الخط میں لکھا متن بھی اگر مشین ریڈ ایبل ہو گا تبھی کوئی ٹرانسلیشن سافٹ ویئر اُسے کسی دوسری زبان میں ترجمہ کرنے کے قابل ہو سکے گا۔ اردو کے رسم الخط کے مشین ریڈ ایبل ہونے کا مطلب اِس کا یونی کوڈ (Unicode) میں لکھا ہونا ہے۔

2.2 ٹرانسلیشن سافٹ ویئر ذریعے اردو ترجمہ: نقلِ حرفی (Transliteration)

متن کے یونی کوڈ میں لکھے ہونے کا ایک فائدہ یہ بھی ہے کہ ترجمے کے عمل کے دوران میں کوئی لفظ اگر مشین کی فہم سے بالاتر ہو رہا ہو تو سافٹ ویئر اُسے نقلِ حرفی کر کے لکھ دیتا/سکتا ہے۔ اِس طرح غیر مانوس یا کمپیوٹر کے لیے عسیر الفہم الفاظ کا ترجمہ بالکل نہ ہونے کی بجائے کچھ نہ کچھ سوا ضرور حاصل ہو جاتا ہے۔

اردو رسم الخط کی رومن حروف میں نقلِ حرفی ایک اہم موضوع ہے۔ آج دنیا بھر میں اردو لکھنے کے لیے رومن رسم الخط بھی پورے زور شور سے استعمال ہو رہا ہے۔ اب تو ایسے لغت آن لائن بھی موجود ہیں جو اردو الفاظ کے معنی رومن حروف میں دیتے ہیں۔ یہ لغات عام لوگوں کے ساتھ ساتھ اخباری نمائندوں کی بڑی ضرورت پوری کرتے ہیں۔ ایک ایسا لغت ملاحظہ کیجیے:

www.websters-online-dictionary.org/translation/Urdu+%2528Transliterated%2529/

مزید دیکھیے: http://www.dictionaryurdu.com/

نقلِ حرفی کا بڑا استعمال متن سے آواز (Text to Speech) میں ہوتا ہے۔ متن کو آواز میں تبدیل کرنے والے بھی سافٹ ویئر اِسی تکنیک سے کام کرتے ہیں۔

2.3 اردو اور پاکستانی زبانوں میں ترجمہ کرنے والے سافٹ ویئر

اردو اور دیگر پاکستانی زبانوں میں مشینی ترجمے کی سہولیات آہستہ آہستہ سامنے آ رہی ہیں۔ پاکستان کا مشینی ترجمے کا پہلا سوا جہد اب انٹرنیٹ پر موجود ہے۔ اِس کا نام PakTranslations.com ہے۔ ایک وقت میں آپ انگریزی کے پانچ ہزار تک الفاظ و فقرات پر مشتمل متن اِس میں اردو ترجمہ حاصل کرنے کے لیے ڈال سکتے ہیں۔ یہ لوگ سندھی میں ترجمہ کی سہولت بھی عنقریب مہیا کر رہے ہیں۔ اِس سائٹ سے آپ پوری پوری ویب سائٹ کو بھی اردو میں ترجمہ کرا سکتے ہیں۔ انگریزی سے اردو اور اردو سے انگریزی کی نقلِ حرفی کی سہولت بھی یہاں مہیا ہے۔ تقریباً تین لاکھ الفاظ پر مشتمل ایک اردو-انگریزی اور انگریزی-اردو لغت بھی اِس سائٹ پر موجود ہے۔

پچھلے صفحات میں جملوں کا ترجمہ کرنے والے سافٹ ویئروں کا ذکر ہوا ہے۔ جہاں تک صرف الفاظ کے ترجمے کی بات ہے، بہت سے ڈکشنری سافٹ ویئر عام طور سے دستیاب ہیں جو ایک زبان کے لفظ کا ترجمہ دوسری زبان میں کر دیتے ہیں۔ یہ سب سافٹ ویئر ٹرانسلیشن میموری (TM) کی بنیاد پر کام کرتے ہیں۔ اردو کی حد تک ایسے کچھ سافٹ ویئروں کا تعارف ملاحظہ کیجیے:

http://www.urduweb.org/mehfil/showthread.php?t=13469

نیز دیکھیے: cleantouch-urdu-dictionary.software.informer.com/

یاد رہے کہ ایسے سافٹ ویئروں کی دو بنیادی اقسام ہیں: آن لائن اور آف لائن۔ جیسا کہ نام سے ظاہر ہے یہ دونوں طرح کے

سافٹ ویئر اپنا اپنا دائرہ کار رکھتے ہیں۔ آف لائن ترجمہ کرنے کی سہولت کے لیے سافٹ ویئر کو اپنے کمپیوٹر پر انسٹال کرنا پڑتا ہے جب کہ آن لائن ترجمے کے لیے ایسی کوئی پابندی نہیں۔ ذیل میں اِن دونوں کی ایک ایک نمائندہ سائٹ دی جا رہی ہے۔

۱۔ آن لائن ترجمے کے لیے: http://www.crulp.org/oud/default.aspx#

۲۔ آف لائن ترجمے کے لیے: http://www.freelang.net/dictionary/urdu.php

2.4 اردو رسم الخط میں اردو مشینی ترجمہ: کرنے کے کام

مشینی ترجمے کے طریقہ ہائے کار کا تفصیلی تعارف پڑھنے سے قارئین نے اندازہ کر لیا ہوگا کہ ہر کامیاب تکنیک میں کارپس تو چاہیے ہی ہے۔ چنانچہ اردو والوں کو سب سے پہلے:

۱۔ کم سے کم ایک بہت بڑا کارپس بنانا ہوگا (صفوان محمد: ۲۰۰۹ء) جو اردو کے روایتی رسم الخط (Indo-Perso-Arabic Script) میں ہو؛

۲۔ اِس کارپس کے متن کو معیاری بنانا ہوگا۔

۳۔ اِس معیاری متن کو مشین ریڈ ایبل اردو یعنی یونی کوڈ میں لکھنا ہوگا؛

۴۔ جب تک یک کارپس ذولسانی نہیں ہوگا (مثلاً اردو-انگریزی یا اردو-عربی وغیرہ، یعنی کسی ایسی زبان کے ساتھ اردو کو جوڑنا جس کا کسی تیسری زبان کے ساتھ ذولسانی کارپس بنا ہوا ہو تا کہ اِس کارپس کی بنیاد پر اردو سے کسی تیسری زبان میں ترجمہ بھی بہ سہولت ہو سکے اور یہ سلسلہ اور زبانوں تک پھیل سکے) تب تک یہ مشینی ترجمے کے لیے کارآمد نہیں ہوگا؛

۵۔ اِس کارپس کو معیاری رومن اردو میں ڈھالنے کی فکر بھی ابھی سے کی جائے کیونکہ پوری دنیا میں بہت سے لوگ اردو صرف اِس لیے نہیں سمجھ پاتے کہ وہ اِس کے رسم الخط سے ناواقف ہیں۔ کسی کو اردو پڑھنے لکھنے سے صرف اِس لیے روک دینا کہ وہ اِس کا رسم الخط نہیں جانتا، بڑی شقاوت ہے۔

3 کمپیوٹر کے ذریعے ترجمہ اور متعلقہ سہولیات — لوگ آج کہاں کھڑے ہیں؟

دنیا ایک عالمی گاؤں بن چکی ہے۔ لوگوں کو ایک دوسرے کے بارے میں جاننے کی جتنی خواہش اور ضرورت آج ہے اب سے پہلے شاید کبھی نہ تھی۔ اِس کی وجوہات سیاسی بھی ہیں اور معاشی بھی۔ ایک دوسرے کے خیالات کو ایک دوسرے کی زبانیں جانے بغیر جانا ممکن نہیں ہوتا۔ ہندوستان میں برطانوی راج نے اپنے قدم یہاں کی علاقائی اور دفتری زبانیں سیکھ کر ہی مضبوط کیے تھے۔ زبانیں سیکھنے اور ایک دوسرے کی زبانوں کو باہم ترجمہ کرنے کا کام آج کمپیوٹر کے ذریعے کیا جا رہا ہے۔ عالمی ترجمہ کار (Universal Translator) جو مدتوں سائنس فکشن کا ایک کھلونا رہا ہے اب تیزی سے حقیقت بن رہا ہے۔ ایک ہاتھ میں اُٹھائی جانے والی ایک زبان سے دوسری زبان میں ترجمہ کرنے کی جیبی مشینیں دن بہ دن بن رہی اور دھڑا دھڑ فروخت ہو رہی ہیں۔

موبائل فونوں پر بھی متن سے آواز (Text to Speech) کی سہولت مہیا ہے۔ بیسیوں سائٹس پر ایسی ٹولز موجود ہیں جو وہ ام

پڑھ کر سنا دیتی ہیں جو آپ نے اپنے موبائل فون کی ایڈریس لسٹ میں محفوظ کیے ہوئے ہوں۔ متن سے آواز کی یہ سہولت بائیں سے دائیں لکھی جانے والی سبھی زبانوں کے ساتھ ساتھ دائیں سے بائیں لکھی جانی والی کچھ زبانوں مثلاً عربی اور عبرانی کے لیے کام کرتی ہے۔ اردو اور فارسی کے حروف تہجی پر مشتمل ایسے ناموں کو جو موبائل فون کی ایڈریس لسٹ میں ہوں، آواز میں منتقل کرنا ابھی شروع ہوا چاہتا ہے۔

مشینی ترجمے کے سلسلے میں اب تک کی سب سے نئی چیز جو ۱۶/فروری ۲۰۱۰ کو بلا قیمت فراہم کی گئی ہے واکساکس (VoxOx) کمپنی کا عالمی ترجمہ کار (Universal Translator) ہے۔ یہ ترجمہ کار انٹرنیٹ پر چیٹنگ، ایس ایم ایس، ای میل اور سماجی میڈیا کے پروگراموں کو اسی لمحے (Real-Time) مختلف زبانوں میں ترجمہ کر دیتا ہے جن میں وہ صارفین کو مطلوب ہوں۔ دونوں صارف اپنی اپنی زبان میں لکھتے ہیں اور یہ سافٹ ویئر اسی لمحے دونوں کو ایک دوسرے کی زبان میں ترجمہ پیش کرتا رہتا ہے۔ یہ سہولت انگریزی اور عربی سمیت بہت سی (خصوصاً یوروپی) زبانوں میں مہیا ہے۔ جلد ہی یہ سہولت کئی لوگوں کو کئی زبانوں میں ترجمے کی فوری فراہمی تک پھیل جائے گی۔

مشینی ترجمے کے ضمن میں تازہ تکنیکی کوششیں یہ ہو رہی ہیں کہ قواعدی اور شماریاتی تکنیکوں کی خوبیوں اور بہتریوں کو جمع کر کے کوئی ایسی دوغلی تکنیک سامنے لائی جائے جس سے ترجمہ کی سہولت عام اور بہتر ہو سکے۔ Google Translator اور Yahoo BabelFish جو اِس وقت دنیا سے مقبول ترین مشینی ترجمہ کار سافٹ ویئر ہیں، اِسی سمت میں آگے بڑھ رہے ہیں۔ امید ہے کہ اِس راستے میں پیش آنے والے مسائل جلد حل ہو جائیں گے کیونکہ اب نہ صرف دنیا کے بڑے دماغ اور رضا کارانہ بنیاد پر کام کرنے والے بے شمار نوجوان اِس عمل میں شریک ہو رہے ہیں بلکہ اِس میں دنیا کے بڑے بڑے کاروباریوں کی دلچسپی بھی پیدا ہو گئی ہے۔ ترجمہ کاری اِس وقت ایک بڑا نفع بخش کاروبار ہے جس سے صرف معاشی ہی نہیں بلکہ عالمی قوتوں کے طویل المیعاد سیاسی فوائد بھی وابستہ ہیں۔ مرتی ہوئی زبانوں اور دنیا کی بڑی زبانوں جنھیں اُن کے بولنے والوں نے ہلکا جان کر خود ہی ڈرر ادیا تھا — جن میں کی سب سے بڑی مثال اردو کی ہے — اب اِن کے کعبوں کو کلیساؤں سے پاسبان مل گئے ہیں۔ آج (۲۷/فروری ۲۰۱۰ء) Google Translator ۱۳۹ زبانوں میں جب کہ Yahoo BabelFish ۷۵ زبانوں میں مفت ترجمے کی سہولت دے رہا ہے۔ حاشیائی عبارت میں اِن زبانوں کے نام اور مزید معلومات ملاحظہ کیجیے۔ اِن سافٹ ویئروں کے ذریعے بیشتر زبانوں کے پورے ویب صفحے اور Blogs ترجمہ ہو جاتے ہیں۔

کاروباری مسابقت کی برکتوں سے آج ایک متن کا ترجمہ بیک وقت ایک سے زیادہ جگہوں سے حاصل کیا جا سکتا ہے۔ آپ اپنے حسب حال بہتر ترجمے کا انتخاب کر سکتے ہیں۔ ترجمے کی سہولت فراہم کرنے والے ادارے اپنی مصنوعات کو بہتر سے بہتر کر رہے ہیں۔ مثلاً ایک سائٹ ForeignWord.com ایک وقت میں ۴۸ مختلف ترجمہ کاروں سے ۳۸ مختلف زبانوں میں ترجمے کی سہولت فراہم کر رہی ہے۔ ملاحظہ کیجیے: www.foreignword.com/Tools/transnow.htm

اِن سافٹ ویئروں کا استعمال بھی آسان ہے۔ آپ متن کو کاپی کر کے ایک مخصوص جگہ پر رکھ دیں، مطلوبہ زبان منتخب کریں، اور Translate کا بٹن دبا دیں۔ مشینی ترجمہ فوراً آپ کے سامنے آ جائے گا۔

آج اِنٹرنیٹ پر ہر ویب براؤزر میں Translate کا بٹن مہیا ہے۔ لہٰذا فی الوقت خواہ یہ نتائج خام ہی کیوں نہ ہو، کسی بھی زبان میں لکھا گیا ویب صفحہ کسی بھی دوسری زبان میں ترجمہ ہونے کا کم سے کم امکان ضرور پیدا ہو گیا ہے جو وقت گزرنے کے ساتھ ساتھ حقیقت میں

بدل رہا ہے۔

ترجمے کی ان سہولیات کے بارے میں یہ بات خاطر نشان رہے کہ یہ صرف آن لائن کام کر سکتے ہیں۔ ترجمے کی آف لائن سہولیات الگ چیز ہیں، جو بالکل محدود اور ایک آدھ زبان تک کے لیے ہوتی ہیں۔ دراصل جب ہم ''کسی بھی زبان سے کسی بھی دوسری زبان'' کی بات کرتے ہیں تو امکانات اتنے زیادہ ہو جاتے ہیں کہ آف لائن ترجمہ ممکن نہیں رہتا۔ ترجمے کی ان سہولیات کو استعمال کرانے کے لیے بہت بڑی بڑی لائبریریاں اور ڈیٹا ذخائر (Repositories) ہیں جو پس منظر میں کام کرتے ہیں؛ یہ از خود یا خودکار طریقوں سے اَپ ڈیٹ ہوتے رہتے ہیں اور ساری دنیا سے ترجمہ کی درخواست کرنے والے لوگوں کو ہمہ وقت دستیاب ہوتے ہیں۔ چونکہ ان کے بغیر ترجمہ ممکن نہیں ہوتا، لہٰذا ان سے استفادہ بھی آن لائن ہوئے بغیر نہیں ہو سکتا۔

ترجمے کی کچھ شکلیں ایسی بھی ہیں جو اَب بن مانگے مہیا ہیں۔ اِس کی سب سے شاندار مثال مترجم ٹوٹکہ (TranslationScreen Tip) ہے۔ آپ کمپیوٹر پر کسی بھی پروگرام مثلاً اِنٹرنیٹ ایکسپلورر، میں کوئی بھی ویب صفحہ کھولیے، اپنے ماؤس کا پوائنٹر کسی بھی لفظ پر لے جا کر ذرا سا ٹھہرائیے یا کھڑا کیجیے، آپ کو پوائنٹر کے سرے پر ایک ڈبے میں اُس لفظ کا ترجمہ (یعنی مترادف لفظ) کتنی ہی زبانوں میں فوراً ملے گا، اور ان میں سے ہر لفظ سے متعلق ڈھیر ساری لسانی معلومات جیسے جنس وغیرہ بھی، اور اِن سب الفاظ کی جملوں میں استعمال کی مثالیں بھی۔ یہ مترجم ٹوٹکہ سہولت صرف آن لائن ہی نہیں بلکہ آف لائن بھی مہیا ہے اور مائیکروسافٹ کے تو سبھی صارف پروگراموں میں ملتی ہے۔ یہ سہولت یقیناً لغت کی بنیاد پر کام کرتی ہے اور اِس میں فرانسیسی، ہسپانوی اور عربی وغیرہ زبانوں میں ترجمہ عام طور سے ملتا ہے۔ مترجم ٹوٹکے نے یہ عالمگیر آسانی پیدا کر دی ہے کہ اب سمجھ نہ آسکنے والی کسی بھی زبان میں بنی ویب سائٹ کا ترجمہ کر کے نیا کرانے کی ضرورت نہیں ہے چونکہ کام چلاؤ ترجمہ جس سے لب لباب سمجھ آسکے، ہر صارف کے سامنے خود ہی آجاتا ہے۔ اِس مترجم ٹوٹکے میں اردو میں ترجمے کی سہولت تو خدا ہی جانے کہ کب آپائے گی۔ ؎

بغض نے پھونک دیا گلشنِ اردو کو حفیظ
آنسوؤں سے ترے یہ باغ ہرا کیا ہوگا

3.1 مشینی ترجمہ کاری کے اطلاقات (Applications)

اوپر کی گئی گفتگو میں جگہ جگہ ایسے سافٹ ویئر پروگراموں کا تعارف دیا گیا ہے جو مختلف انداز میں مشینی ترجمہ کاری کی سہولت فراہم کر رہے ہیں۔ یہ پروگرام زیادہ تر آن لائن کام کرتے ہیں۔ اِس سلسلے میں کچھ ضروری معلومات کو یہاں یکجا کیا جا رہا ہے:

۱۔ Asia Online ایک ایسا سافٹ ویئر ہے جو ایک خاص انداز میں بنائے ہوئے ترجمہ انجن کے ذریعے انسانی زبان کے قریب ترین (Near-human quality) ترجمہ فراہم کرنے کا دعویٰ کرتا ہے۔

۲۔ ہندی سے پنجابی ترجمہ کار: یہ ترجمہ کار بلاواسطہ (Direct) طریقے سے ترجمہ کرتا ہے۔ یہ ہندی ویب سائٹوں کو پنجابی میں ترجمہ کرنے کی سہولت بھی دیتا ہے اور ہندی میں لکھی ای میل کو پنجابی میں ترجمہ کر کے بھیجنے کی صلاحیت بھی رکھتا ہے۔

۳۔ WorldLingo قواعدی اور شماریاتی ترجمہ کار طریقوں کو بیک وقت استعمال کرتا ہے۔ اِس پروگرام کی کارگزاری اور کاروباری ساکھ کے بارے میں یہ کہنا کافی ہے کہ یہ مائیکروسافٹ کا کاروباری ساتھی ہے۔

۴۔ AppTek نے ۲۰۰۹ء میں دوغلا مشینی ترجمہ سسٹم پیش کیا ہے وغیرہ۔

اِن کے علاوہ بڑے سافٹ ویئر مثلاً Trados، Wordfast، Swordfish اور CATALYST Alchemy وغیرہ زیادہ تر مائیکروسافٹ ہی کے پلیٹ فارم پر کام کرتے ہیں۔ ترجمہ کار سافٹ ویئر زیادہ تر مفت مہیا ہیں مثلاً ForeignDesk۔ یہ بات اظہر من الشمس ہے کہ کوئی بھی مشینی ترجمہ کار سافٹ ویئر ''کامل خودکار اعلیٰ درجے کا ترجمہ کار'' (FAHQT) نہیں ہوتا جو ہر آزاد متن (Unrestricted text) کا معقول ترجمہ کر سکے۔ اگر متن مخصوص اور کسی ضابطے کا ہو تو مشینی ترجمے کا معیار بہت بہتر ہو جاتا ہے۔ اپنی محدودات کے باوجود مشینی ترجمہ کار پروگراموں کی مانگ دنیا بھر میں ہے۔ اِن پروگراموں کا سب سے بڑا ہندھا لگا گاہک شاید یوروپین کمیشن (European Commission) ہے۔

پچھلے کچھ عرصے میں دہشت گردی کی کارروائیوں میں اضافہ ہونے کی وجہ سے امریکی فوجی ادارے انسانی زبانوں کے ترجمے پر بہت زیادہ رقوم خرچ کر رہے ہیں۔ In-Q-Tel ایک ایسی ہی مہم کا نام ہے جس میں امریکی خفیہ ایجنسیاں زیادہ تر اپنے خرچ پر پرائیویٹ سیکٹر کو بیچ کا بچولیا بنا کر اچھی تکنیکی صلاحیتیں رکھنے والی کمپنیوں سے یہ تحقیقات کروا رہی ہیں۔ اپنے مخصوص عزائم اور جغرافیائی ضروریات کی وجہ سے اِس وقت فوجی جتھا عربی، پشتو اور دری زبانوں کے مشینی ترجمے اور تحقیق میں زیادہ دلچسپی لے رہی ہے۔ اِسی طرح اردو کو مشین ریڈ ایبل کرنا اور اردو مشینی ترجمہ بھی علمی سے زیادہ ایک سیاسی ضرورت ہے جس سے عالمی طاقتوں کے مفادات وابستہ ہیں۔

پچھلے چند سالوں میں اِنٹرنیٹ کے دنیا بھر کے سماج میں پھیل جانے سے آنے والی کاروباری اُٹھان نے بھی مشینی ترجمے کے مواقع میں یکایک اضافہ کر دیا ہے۔ Facebook وغیرہ جیسے سماجی نیٹ ورک پروگراموں کے ساتھ ساتھ فوری پیغامات (Instant Messaging) بھیجنے والے سافٹ ویئر مثلاً GoogleTalk اور MSN Messenger وغیرہ اپنے استعمال کنندگان کو کئی زبانوں میں متن لکھنے کی سہولت دے رہے ہیں۔ موبائل مصنوعات مثلاً موبائل فون، سفری کمپیوٹر (Laptop) اور ڈیجیٹل ڈائریوں وغیرہ پر یہ سہولتیں عام ہیں۔ اِن مصنوعات کی ساخت ایسی بنائی جا رہی ہے کہ باہم مختلف زبانیں بولنے والے کاروباری شراکت دار اور گاہک بغیر مترجم کے صرف اِنہی مشینوں کے استعمال سے ایک دوسرے کی بات سمجھ سکیں۔

3.2 مشینی ترجمہ کاری کا مستقبل

مشینی ترجمہ کاری اپنے ابتدائی اور آزمائشی مراحل عبور کر چکی ہے اور اب سرعت سے کامیابیوں کی منزلیں طے کر رہی ہے۔ یوں محسوس ہوتا ہے کہ اب عام زندگی میں بھی ترجمے کا کام ہر جگہ مشینوں ہی کی معاونت سے ہوا کرے گا۔ جلد ہی ہر ویب صفحے کے ساتھ مشینی ترجمہ کار منسلک ہوگا۔ اور پھر اِس ترجمہ شدہ مواد کی متن سے آواز اور آواز سے مطلوبہ فارمیٹ میں متن میں تبدیلی کی سہولت بھی کھڑے پیر میسر ہو جائے گی۔ یہ بھی ممکن ہے کہ آئندہ دنوں میں دفتروں کے اندر مشینی ترجمہ کار فوٹو کاپی مشینوں، کیمروں، سکینروں اور کارڈ ریڈروں کے ساتھ منسلک ہو کر ترجمے کی خدمات انجام دے رہے ہوں۔

مشینی ترجمہ کاری کی معراج یہی ہوگی کہ زبانوں کا آڑا کھوڑ (Barrier) بالکل ختم ہو جائے گا: ایک آدمی دنیا کی کسی بھی زبان میں اظہارِ خیال کر رہا ہو تو سننے والے اُسی لمحے اُسے اپنی اپنی (یا اپنی مطلوبہ) زبانوں میں ترجمہ ہو کر سن اور سمجھ رہے ہوں گے، اور یہ متن اُسی لمحے لکھا اور شائع بھی کیا جا رہا ہوگا۔

یہ مشینی ترجمہ کاری کی تصویر کا پرلا رخ ہے؛ اِدھر لا رخ یہ ہوگا کہ انسانی دماغوں کے اندر ایسی Chip لگانے اور اُسے حسبِ ضرورت تبدیل کرنے صورت پیدا کر لی جائے گی جو اِس پورے نام جہام کو کنٹرول کر رہی ہو۔ یہ بھی ممکن ہے کہ جنیاتی انجینئرنگ (Genetic Engineering) کے ذریعے جہاں انسان آج انتخابِ جنس، مختلف اعضائے جسمانی کے رنگ اور ذہانت وغیرہ جیسے قدرتی عوامل میں تصرف کر چکا ہے وہیں مادری طور پر ایک سے زیادہ زبانیں بلوانے اور جنوانے کی قدرت بھی حاصل کر لے۔ معلوم تاریخِ انسانی میں یہ واقعہ رونما ہو چکا ہے کہ ایک صبح کچھ لوگ اُٹھے تو وہ اپنی مادری زبانوں کو بھول کر مختلف قوموں کی زبانیں بول رہے تھے۔ یہ محیرالعقل واقعہ تب پیش آیا جب حضرت عیسیٰ علیہ السلام نے اپنے کچھ حواریوں کو مختلف زبانیں بولنے والی قوموں میں تبلیغِ دین کے لیے بھیجنا چاہا تھا اور انھوں نے اُن کی زبانیں نہ جاننے کا عذر کیا تھا۔ قدرت کا یہ تماشہ (Phenomenon) دوبارہ بھی — اور پہلے سے کہیں زیادہ شدت اور امکانات کے ساتھ — رونما ہو سکتا ہے۔

آدم وحوا ایک ہی زبان بولتے تھے قطعِ نظر اِس سے کہ اُن کی زبان کیا تھی، اُن کی اولاد ہزاروں زبانیں بول رہی ہے ۔

یہ ڈرامہ دکھائے گا کیا سین
پردہ اُٹھنے کی منتظر ہے نگاہ

4: خاتمہ

کمپیوٹر کی معاونت سے ایک زبان سے دوسری زبان میں ترجمہ کرنا آج کی تحقیقات وترقیات میں ایک اہم مقام رکھتا ہے اور جس پر بہت انسانی اور مادی وسائل خرچ کیے جا رہے ہیں۔ یہ کام مشینی ترجمے کی مختلف تکنیکوں کو بروئے کار لاتے ہوئے کیا جاتا ہے۔ اِس مقالے میں مشینی ترجمہ کی تاریخ، موجودہ صورتِ حال اور اِس کے مستقبل کے امکانات کا ایک مطالعہ کمپیوٹر اور اطلاعیات کی زبان میں پیش کیا گیا ہے۔ دنیا کی کسی بھی زبان سے اردو میں اور اردو سے دنیا کی کسی بھی زبان میں مشینی ترجمہ اب بارہ پتھر دور کی بات نہیں رہی۔ یہ فاصلہ اب روشنی کی رفتار سے طے ہو رہا ہے۔

☆☆☆

تحریر: ۲۷/فروری ۲۰۱۰ء مطابق ۱۲/ ربیع الاولی ۱۴۳۱ھ

مزید مطالعہ:

اِس مقالے کے قارئین سے درخواست ہے وہ مندرجہ ذیل مقالات کو بھی توجہ سے پڑھ ڈالیں:

۱۔ بخاری، سید ذوالکفل؛ ظہیر احمد، ڈاکٹر؛ صفوان محمد چوہان، ڈاکٹر حافظ؛ ۲۰۰۹ء اردو کارپس: تکنیکی تعارف، اہمیت،

ضرورت اور دائرہ و لائحۂ عمل۔مشمولہ:جرنل آف ریسرچ، بہاء الدین زکریا یونیورسٹی ملتان،شمارہ-۱۴۔

۲۔ صفوان محمد چوہان،ڈاکٹر حافظ؛۲۰۱۰ء،اردو اطلاعیات: آج اور کل،مشمولہ:سہ ماہی اثبات،شمارہ:۵-۴۔

۳۔ صفوان محمد چوہان،ڈاکٹر حافظ؛۲۰۰۷ء،ترویج اردو کی ایک فوری ضرورت: اردو رسم الخط میں انگریزی-اردو لغات کی آن لائن فراہمی۔مشمولہ:اورینٹل کالج میگزین۔جلد-۸۳،عدد:۲-۱۔

حواشی وتعلیقات

حوالہ جات:

بخاری،سید ذوالکفل ظہیر احمد،ڈاکٹر؛صفوان محمد چوہان،ڈاکٹر حافظ؛۲۰۰۹ء،اردو کارپس: تکنیکی تعارف، اہمیت، ضرورت اور دائرہ و لائحۂ عمل۔مشمولہ:جرنل آف ریسرچ، بہاء الدین زکریا یونیورسٹی ملتان،شمارہ-۱۴۔

حواشی

۱۔ ملاحظہ کیجیے: www.hutchinsweb.me.uk/MTNI-14-1996.pdf

۲۔ اردو افعال،سونیا چرنی کووا(Sonia Charni Kova)،ترقی اردو بیورو،نئی دہلی، اپریل،جون ۱۹۸۹ء،ص-۲۹۔

۳۔ ملاحظہ کیجیے: http://crulp.org/

۴۔ ملاحظہ کیجیے:اردو کارپس: تکنیکی تعارف،اہمیت، ضرورت اور دائرہ و لائحۂ عمل،مشمولہ جرنل آف ریسرچ، بہاءالدین زکریا یونیورسٹی ملتان،شمارہ-۱۴،۲۰۰۸۔ص-۱۳۵۔

۵۔ ملاحظہ کیجیے: http://www.voxox.com/home.php

۶۔ Google Translator کی مشینی ترجمہ کاری سہولت مندرجۂ ذیل زبانوں کے لیے دستیاب ہے۔ملاحظہ کیجیے:
www.google.com/language_tools?hl=EN

•Afrikaans	•Esperanto	•Javanese	•Nepali	•Somali
•Akan	•Estonian	•Kannada	•Norwegian	•Spanish
•Albanian	•Faroese	•Kazakh	•Norwegian (Nynorsk)	•Sundanese
•Amharic	•Filipino	•Kinyarwanda	•Occitan	•Swahili
•Arabic	•Finnish	•Kirundi	•Oriya	•Swedish
•Armenian	•French	•Klingon	•Oromo	•Tajik
•Azerbaijani	•Frisian	•Korean	•Pashto	•Tamil
•Basque	•Galician	•Kurdish	•Persian	•Tatar
•Belarusian	•Georgian	•Kyrgyz	•Pirate	•Telugu
•Bengali	•German	•Laothian	•Polish	•Thai
•Bihari	•Greek	•Latin	•Portuguese (Brazil)	•Tigrinya
•Bork	•Guarani	•Latvian	•Portuguese (Portugal)	•Tonga
•Bosnian	•Gujarati	•Lingala	•Punjabi	•Turkish
•Breton	•Hacker	•Lithuanian	•Quechua	•Turkmen
•Bulgarian	•Hausa	•Luganda	•Romanian	•Twi
•Cambodian	•Hawaiian	•Macedonian	•Romansh	•Uighur
•Catalan	•Hebrew	•Malagasy	•Russian	•Ukrainian
•Chinese (Simp)	•Hindi	•Malay	•Scots Gaelic	•Urdu
•Chinese (Trad)	•Hungarian	•Malayalam	•Serbian	•Uzbek
•Corsican	•Icelandic	•Maltese	•Serbo-Croatian	•Vietnamese
•Croatian	•Igbo	•Maori	•Sesotho	•Welsh
•Czech	•Indonesian	•Marathi	•Shona	•Xhosa
•Danish	•Interlingua	•Mauritian Creole	•Sindhi	•Yiddish
•Dutch	•Irish	•Moldavian	•Sinhalese	•Yoruba
•Elmer Fudd	•Italian	•Mongolian	•Slovak	•Zulu
•English	•Japanese	•Montenegrin	•Slovenian	

۷۔ Yahoo BabelFish کی مشینی ترجمہ کاری کی سہولت کے لیے ملاحظہ کیجیے: babelfish.yahoo.com اِس سہولت کا تعارف اِن الفاظ میں کرایا گیا ہے: Language translator used for translating text or web pages from English and other languages to French, German, Greek, Chinese, Spanish, and others.

۸۔ ملاحظہ کیجیے: http://en.wikipedia.org/wiki/Asia_Online

۹۔ ملاحظہ کیجیے: http://h2p.learnpunjabi.org/

۱۰۔ ملاحظہ کیجیے: http://en.wikipedia.org/wiki/Worldlingo اگرچہ یہ صفحہ ۵/دسمبر ۲۰۰۹ کو وکی پیڈیا سے ہٹا دیا گیا ہے۔

۱۱۔ ملاحظہ کیجیے: http://en.wikipedia.org/wiki/Apptek

۱۲۔ ملاحظہ کیجیے: http://sourceforge.net/projects/foreigndesk/

تشکر (Acknowledgement):

۱۔ یہ مقالہ لکھنے میں جناب وصی اللہ کھوکھر نے میری بہت مدد کی ہے اور بے حد اہم معلومات فراہم کی ہیں۔ اُنھوں نے اِس مقالے کو چھپنے سے پہلے جانچا بھی ہے۔ رسمی شکریے کے کوئی الفاظ اُن کی خدمات کا بدل نہیں ہو سکتے۔ جناب وصی اللہ کھوکھر اردو مشینی ترجمے میں پاکستان کے پرانے لوگوں میں سے ہیں اور کمپیوٹر اور اطلاعیاتی ٹیکنالوجی کی پندرہ سے زیادہ کتابوں کے مصنف/ مترجم ہیں۔ وہ مائیکروسافٹ کے ساتھ اردو ترجمے کے کئی ایک پراجیکٹ کر چکے ہیں۔ کمپیوٹر سائنس کی اصطلاحات کو اردو-انگریزی میں جس سہولت اور تخلیقی صلاحیت سے وہ استعمال کرتے ہیں، صرف پاکستان ہی نہیں بلکہ پوری دنیا میں بھی ایسے کا ئیاں لوگ کم ہی ہوں گے۔

۲۔ اِس مقالے کی تیاری کے دوران میں حوالے کی بہت سی باتوں کے ضمن میں ڈاکٹر شمس الرحمٰن فاروقی، ڈاکٹر گوپی چند نارنگ، ڈاکٹر خواجہ محمد زکریا اور ڈاکٹر خورشید رضوی صاحب کو بار بار زحمت دی گئی۔ ہر چار حضرات کا شکریہ واجب ہے۔

۳۔ یہ مقالہ لکھنے کے لیے www.wikipedia.org کے مختلف مقالات سے آزادانہ استفادہ کیا گیا ہے۔

فہرستِ اسنادِ محولہ

مآخذ

الف: کتابیات

۱۔ قرآنِ پاک۔ The Message of THE QURAN by Muhammad Asad, Dar al-Andalus Limited, 3 Library Ramp, Gibralter. 1980.

۲۔ اقبال، شاعرِ مشرق ڈاکٹر علامہ سر محمد، کلیاتِ اقبال، پانچواں ایڈیشن، اقبال اکادمی پاکستان، لاہور ۲۰۰۰ء

۳۔ حفیظ جالندھری، ابوالاثر، کلیاتِ حفیظ، پہلا ایڈیشن، مرتبہ ڈاکٹر خواجہ محمد زکریا، الحمد پبلی کیشنز، لاہور دسمبر ۲۰۰۵ء

۴۔ عابد صدیق، پانی میں ماہتاب، دوسرا ایڈیشن، الحمد پبلی کیشنز، لاہور ۲۰۰۶ء

۵۔ غالب، مرزا اسد اللہ خاں، دیوانِ غالب۔ غالب انسٹی ٹیوٹ، نئی دہلی۔ فروری ۱۹۸۶ء

۶۔ Europe Speaks Arabic by Dr V Abdul Rahim, Institute of the Language of the Quran Inc, Toranto, Canada, 2008.

ب: رسائل اور تحقیقی جرائد

۱۔اردو اطلاعیات آج اور کل ۔مشمولہ: سہ ماہی اردو نامہ، مجلس زبان دفتری حکومت پنجاب ۔شمارہ اکتوبر ۲۰۰۸ء تا مارچ ۲۰۰۹ء ۔ص ص ۱۶۳ تا ۱۷۳۔

۲۔اردو اور دنیا کی بڑی زبانوں کی شماریات ۔مشمولہ: اردو سائنس میگزین ۔شمارہ-۴، ۲۰۰۹ء ۔ص ص ۱۹ تا ۲۳۔

۳۔اردو رسم الخط میں انگریزی-اردو لغات کی آن لائن فراہمی ۔مشمولہ: اردو سائنس میگزین ۔شمارہ-۴، ۲۰۰۷ء ۔ص ص ۳۲ تا ۳۷ اور مشمولہ: اورینٹل کالج میگزین ۔ جلد-۸۳، عدد: ۲-۱، ۲۰۰۷ء ۔ص ص ۲۴۳ تا ۲۶۴ ۔

۴۔اردو لغت (تاریخی اصول پر) بدلتے لسانی تناظر میں چند تجاویز ۔مشمولہ: جرنل آف ریسرچ، بہاء الدین زکریا یونیورسٹی ملتان ۔شمارہ-۱۲ (۲۰۰۷ء) ۔ص ص ۲۷۵ تا ۲۸۶۔

۵۔مشین ریڈایبل اردو رسم الخط: حروف کی کشتیاں، اعراب، نقطے، شوشے اور کششیں ۔مشمولہ: اردو سائنس میگزین ۔شمارہ-۳، ۲۰۰۸ء ۔ص ص ۲۹ تا ۳۳۔

ج: انٹرنیٹ سائٹس (چند منتخب سائٹس)

1. http://www.hutchinsweb.me.uk/Nutshell-2005.pdf
2. http://www.isi.edu/natural-language/projects/rewrite/mtsummit03.pdf
3. http://www.machinetranslations.org

تکنیکی مشاورت

۱۔ڈاکٹر خواجہ محمد زکریا، سابق پرنسپل، اورینٹل کالج، جامعۂ پنجاب، لاہور [اردو و انگریزی]

۲۔ڈاکٹر سید خورشید حسن رضوی، لاہور [اردو و عربی]

۳۔خواجہ غلام ربانی مجال، ۳۸-گلستان کالونی، لین نمبر-۲، نیشنل پارک روڈ، راولپنڈی [اردو و انگریزی]

۴۔راؤ صفدر رشید، ڈیٹابیس سپروائزر، مرکزِ فضیلت برائے اردو اطلاعیات، مقتدرہ قومی زبان، اسلام آباد [اردو مشینی ترجمہ]

۵۔وصی اللہ کھوکھر، ایم جی ایچ سالوشنز، کاسوکے [اردو مشینی ترجمہ]

۶۔ڈاکٹر پرویز احمد Language Technologies Research Centre, International Institute of Information Technology, Gachibowli, Hyderabad. [اردو مشینی ترجمہ]