



1359- واحد شهید



جمهوری اسلامی ایران



دانشگاه جامع علمی-کاربردی

دانشگاه جامع علمی- کاربردی

مرکز آموزش عالی علمی - کاربردی جهاد دانشگاهی مشهد

پایان نامه کارشناسی رشته مهندسی فناوری اطلاعات
گرایش برنامه نویسی تحت وب

استخراج الگوهای انسجامی بین یک متن و ترجمه آن
توسط رویکردهای ترجمه آماری ماشینی

استاد راهنما:

جناب آقای دکتر محمد عبداللهی

استاد مشاور:

جناب آقای مهندس نیک فرجام

نگارندگان:

آقای سهیل افروزه، خانم ها مرضیه آتون، زهرا امیدزاده و محبوبه خوشدل

تابستان 1396

مشکر و قدردانی

مَنْ عَلَّمَنِي حَرْفًا فَقَدْ صَيَّرَنِي عَبْدًا «امام علی علیه السلام»

همه‌ی پیشرفت‌های معنوی و مادی بشر ریشه در تربیت افراد دارد که از منبع زلال ایمان، معرفت، تعهد و ایثار معلم سیراب می‌شوند. از این منظر کشور و امدار همه‌ی استادانی است که بحق چراغ هدایت بشریت را به تاسی از پیامبر اعظم (ص) روشن و فروزان نگاه داشته‌اند. فرصت را معتنم شمرده، در کمال ادب و ارادت این تقدیر را تقدیم محضر عالمانه و اندیشمندان‌هی استاد بزرگوار جناب آقای دکتر عبداللہی می‌نمایم. از حضرت علیم آرزو داریم همواره در کسوت شریف استادی با کلام نافذ و عزم راسخ و با تولید علم، دانش و پژوهش جهت آینده‌سازان این مرز و بوم با سرفرازی برای نیل به زندگی سعادتمندان‌هی دنیوی و رستگاری اخروی کوشا و پویا باشید.

فهرست مطالب

| | |
|-----------------------------|--------------------------------|
| 6 | چکیده |
| 7 | مقدمه |
| فصل اول کلیات | |
| 10 | 1-1- مقدمه |
| 10 | 2-1- اهمیت و ضرورت تحقیق |
| 11 | 3-1- اهداف مشخص تحقیق |
| 11 | 4-1- پرسش‌های تحقیق |
| 11 | 5-1- تعاریف واژه‌ها و اصطلاحات |
| 11 | 1-5-1- کلمه |
| 11 | 2-5-1- جمله |
| 12 | 3-5-1- عبارت |
| 12 | 4-5-1- اصطلاح |
| 13 | 5-5-1- ترجمه |
| 14 | 6-5-1- پیشینه ترجمه در ایران |
| 15 | 7-5-1- ترجمه ماشینی |
| 15 | 8-5-1- انسجام متن |
| 16 | 9-5-1- فرآیند ترجمه |
| 16 | 10-5-1- روش‌ها |
| ERROR! BOOKMARK NOT DEFINED | - روش قاعده‌مند |
| ERROR! BOOKMARK NOT DEFINED | - روش بین‌زبانی |
| ERROR! BOOKMARK NOT DEFINED | - روش مبتنی بر فرهنگ لغت |
| ERROR! BOOKMARK NOT DEFINED | - روش مبتنی بر مثال |
| ERROR! BOOKMARK NOT DEFINED | - ترجمه ماشینی پیوندی |
| ERROR! BOOKMARK NOT DEFINED | رمزگشایی |
| ERROR! BOOKMARK NOT DEFINED | روش آماری |

| | |
|----------------------------------|--|
| ERROR! BOOKMARK NOT DEFINED..... | 11-5-1- مسئله‌های اصلی |
| ERROR! BOOKMARK NOT DEFINED..... | ابهام زدایی |
| ERROR! BOOKMARK NOT DEFINED..... | گفتارهای غیراستاندارد |
| ERROR! BOOKMARK NOT DEFINED..... | واحدهای اسمی |
| ERROR! BOOKMARK NOT DEFINED..... | وجود اصطلاحات و عبارات معنی دار خاص در هر زبان |
| ERROR! BOOKMARK NOT DEFINED..... | فناوری تعریف شده در تحقیق |

فصل دوم: پژوهش‌های پیشین

| | |
|----------------------------------|---|
| ERROR! BOOKMARK NOT DEFINED..... | 1-2- مقدمه |
| ERROR! BOOKMARK NOT DEFINED..... | 2-2- مختصر اشاره به فن آوری های قبلی |
| ERROR! BOOKMARK NOT DEFINED..... | 1-2-2- قبل از رایانه |
| ERROR! BOOKMARK NOT DEFINED..... | 2-2-2- پیشگامان، ۱۹۴۷ تا ۱۹۵۴ |
| ERROR! BOOKMARK NOT DEFINED..... | 3-2-2- دهه‌ی خوش‌بینی ۱۹۵۴-۱۹۶۶ |
| ERROR! BOOKMARK NOT DEFINED..... | 4-2-2- نتیجه‌ی گزارش ALPAC، ۱۹۶۶ تا ۱۹۸۰ |
| ERROR! BOOKMARK NOT DEFINED..... | 5-2-2- دهه‌ی ۱۹۸۰ |
| ERROR! BOOKMARK NOT DEFINED..... | 6-2-2- اوایل دهه ۱۹۹۰ |
| ERROR! BOOKMARK NOT DEFINED..... | 7-2-2- اواخر دهه ۱۹۹۰ و اوایل دهه ۲۰۰۰ |
| ERROR! BOOKMARK NOT DEFINED..... | 8-2-2- روش‌ها و مدل‌هایی که تا کنون بررسی شده‌اند |
| ERROR! BOOKMARK NOT DEFINED..... | بخش بندی |
| ERROR! BOOKMARK NOT DEFINED..... | توزیع کلمات |
| ERROR! BOOKMARK NOT DEFINED..... | اجزای سخن |
| ERROR! BOOKMARK NOT DEFINED..... | مورفولوژی با ریخت‌شناسی |
| ERROR! BOOKMARK NOT DEFINED..... | تخمین توزیع احتمال |
| ERROR! BOOKMARK NOT DEFINED..... | توزیع‌های احتمال رایج |
| ERROR! BOOKMARK NOT DEFINED..... | روش‌های مبتنی بر کلمه |
| ERROR! BOOKMARK NOT DEFINED..... | ترجمه‌ی ماشینی توسط ترجمه‌ی لغات |
| ERROR! BOOKMARK NOT DEFINED..... | مدل IBM |

فصل سوم: روشها و الگوریتم‌های پیشنهادی

| | |
|----------------------------------|--|
| ERROR! BOOKMARK NOT DEFINED..... | <u>1-3-1</u> مقدمه |
| ERROR! BOOKMARK NOT DEFINED..... | <u>2-3-2</u> روش پیشنهادی در تحقیق |
| ERROR! BOOKMARK NOT DEFINED..... | <u>1-2-3-1</u> مدل‌های آماری هم‌ترازی کلمات (IBMS) |
| ERROR! BOOKMARK NOT DEFINED..... | <u>2-2-3-2</u> مدل N-GRAM |
| ERROR! BOOKMARK NOT DEFINED..... | استخراج مدل N-GRAM |
| ERROR! BOOKMARK NOT DEFINED..... | هموارسازی |
| ERROR! BOOKMARK NOT DEFINED..... | تخفیف |
| ERROR! BOOKMARK NOT DEFINED..... | <u>3-2-3-3</u> مساله کدگشایی و الگوریتم ویتربی (VITERBI ALGORITHM) |
| DEFINED. | |
| ERROR! BOOKMARK NOT DEFINED..... | <u>4-2-3-4</u> مدل مخفی مارکوف (HMM) |
| ERROR! BOOKMARK NOT DEFINED..... | <u>5-2-3-5</u> قضیه بیز |
| ERROR! BOOKMARK NOT DEFINED..... | یادگیری ماشینی به کمک نظریه بیز |
| ERROR! BOOKMARK NOT DEFINED..... | تئوری بیز در یادگیری ماشین |
| ERROR! BOOKMARK NOT DEFINED..... | تعریف مفاهیم اولیه |
| ERROR! BOOKMARK NOT DEFINED..... | روشهای یادگیری بیزی ماشینی |
| ERROR! BOOKMARK NOT DEFINED..... | <u>5-2-3-5</u> ترجمه بر اساس شبکه‌های عصبی چیست؟ |
| ERROR! BOOKMARK NOT DEFINED..... | در شبکه عصبی و سطح بالا کار ترجمه ها در دو مرحله |

فصل چهارم: نتیجه گیری

| | |
|----------------------------------|--------------------------|
| 49..... | <u>1-4-1</u> مقدمه |
| 49..... | <u>2-4-2</u> نتیجه گیری |
| 50..... | <u>3-4-3</u> خلاصه |
| ERROR! BOOKMARK NOT DEFINED..... | مباحث تعریف شده در تحقیق |
| ERROR! BOOKMARK NOT DEFINED..... | مدل HMM |

فصل پنجم: منابع

| | |
|----------------------------------|-------|
| ERROR! BOOKMARK NOT DEFINED..... | مراجع |
|----------------------------------|-------|

چکیده

ترجمه ماشینی زیرشاخه‌ای از زبان شناسی محاسباتی بوده که به ترجمه یک متن از زبانی به زبان دیگر، توسط ماشین و بدون دخالت انسان و با استفاده از الگوهای آماری، دانش زبان شناسی و الگوریتم‌های کامپیوتری می‌پردازد. ترجمه ماشینی نه تنها دارای جاذبه و کشش علمی بوده، بلکه از دیدگاه اقتصادی و دیگر ضرورت‌های اجتماعی دارای اهمیت بوده و تحقیق بر روی آن را کاملاً توجیه می‌کند. یکی از دلایل استفاده از ترجمه ماشینی، افزایش سرعت و کاهش هزینه است. اما مهمترین ویژگی در یک متن ترجمه شده انسجام ترجمه خروجی است. انسجام در یک متن همه آن روابط معنایی است که میان عناصر سازنده آن متن برقرار می‌شوند و تشکیل سلسله‌ای از کلام را می‌دهد، بنابراین «آنچه متن را یکپارچه و از نوشته‌های پراکنده متمایز می‌سازد، انسجام و پیوستگی آن است. اتصال جمله‌ها یا روابط عوامل آن با یکدیگر هنگامی پدید می‌آید که تعبیر عاملی در یک بخش از کلام به عاملی دیگر در همان کلام وابسته باشد. اینجاست که رابطه انسجامی شکل می‌گیرد و این دو عامل سبب شکل‌گیری متن می‌شوند.

ترجمه ماشینی علی‌رغم رشد فناوری‌ها هنوز هم نقص‌های فراوانی دارد. اصولاً چون کامپیوترها مانند انسان هوشمند نیستند، ترجمه‌ای هم که توسط آنها انجام شود، ترجمه‌ی کاملی نخواهد بود. نرم‌افزارهای مترجم، در بهترین حالت، ترجمه را با دقتی حدود 70٪ انجام می‌دهند. این مشکلات در زبان فارسی بسیار بیشتر بوده و آن هم به دلیل وجود ابهام‌های موجود در جملات آن بوده که برای ماشین غیر قابل تشخیص است. به عنوان مثال هنگامی که در ترجمه، ماشین به عبارت یا لغتی برخورد می‌کند که دارای تعداد زیادی از معانی است، نمی‌تواند مناسب‌ترین لغت را انتخاب نماید یا جای کلمات را در داخل جمله عوض کند و خروجی آن معمولاً بیانی غیرطبیعی یا اشتباه خواهد بود و یا حتی نمی‌تواند آرایه‌های ادبی مثل استعاره و تشبیه و... را به طور کامل رعایت کند بنابراین حس نویسنده را به خواننده نمی‌رساند و استفاده از ماشین، ترجمه‌ی متن را دشوار می‌نماید.

در این تحقیق به بررسی روش‌های ترجمه آماری ماشینی پرداخته‌ایم. سعی بر آن است که الگوهای انسجامی بین یک متن و ترجمه آن را توسط روش‌های ترجمه آماری ماشینی استخراج کرده و راهکارهایی برای ارزیابی و بهبود متن ترجمه شده ارائه نماییم.

کلمات کلیدی: ترجمه آماری ماشینی، الگوهای آماری، دانش زبان شناسی، الگوریتم‌های کامپیوتری، انسجام متن ترجمه

شده، مدل‌های زبانی، الگوهای انسجامی

مقدمه:

از پنجاه سال پیش تاکنون، ترجمه ماشینی شاهد پیشرفت‌های شایان توجهی بوده است. بطور کلی روند ترجمه ماشینی به این صورت است که ابتدا کامپیوتر متن مبدأ را به واحدهای کوچکتر، یعنی کلمه یا عبارت، تقسیم می‌کند. سپس آنها را از لحاظ نحوی و معنایی تجزیه و تحلیل و رمزگشایی می‌کند. پس از آن، بر اساس اصول و قواعد ساختاری و معنایی زبان مقصد، معادل متن موردنظر را می‌سازد. اساس کار نسل‌های اول ترجمه ماشینی نیز همین شیوه بوده و معادل واژگان به‌طور مستقیم از واژه‌نامه اتخاذ می‌شده است. از آنجایی که حاصل کار این نوع ترجمه دور از معنای واقعی بود، مورد توجه پیشگامان این فناوری و کاربران قرار نگرفت و دیری نپایید که جای خود را به نسل بعدی داد.

سیستم‌های نسل دوم ترجمه ماشینی از یک زبان واسطه بهره می‌بردند. در این روش، متن مبدأ ابتدا از نظر دستوری و معنایی توسط کامپیوتر بررسی و پس از آن، به یک زبان واسطه مستقل برگردانده می‌شد. در نهایت، معادل کلمات از واژه‌نامه انتخاب و بر اساس قواعد زبان مقصد در جمله قرار داده می‌شوند. به دلیل غیرطبیعی بودن نتیجه کار، روش بین‌زبانی نیز نتوانست رضایت کاربران را جلب کند و با ظهور و گسترش اینترنت، نسل نوین ترجمه ماشینی پا به عرصه فناوری گذاشت.

در طراحی جدید ماشین‌های ترجمه از تکنیک‌های مختلفی شامل روش‌های آماری، مقایسه‌ای، تطبیقی و مبتنی بر مثال برای تبدیل زبان مبدأ به زبان مقصد استفاده شده است. اساس کار برنامه‌های نوین ترجمه آنلاین، بهره‌گیری از ذخایر عظیم پیکره‌های زبانی در اینترنت است. منظور از پیکره زبانی، نمونه‌های ترجمه شده، مثال‌ها و متون معادل موجود از متن مبدأ به زبان مقصد است. کامپیوتر پس از تجزیه و تحلیل متن مبدأ، در حین معادل‌گزینی به این پیکره‌های زبانی آنلاین رجوع کرده و ترجمه خود را با ترجمه‌های مشابه مقایسه می‌کند. سپس با استفاده از روش آماری، حاصل‌نهایی کار را با مناسب‌ترین و رایج‌ترین معادل‌ها مطابقت می‌دهد.

تصور می‌شود که متأسفانه برخی به آن معتقدند این است که ماشین‌های مترجم آمده‌اند تا شغل مترجمی را از صحنه بردارند. بر خلاف این تصور، فناوری ترجمه ماشینی برای جایگزین کردن ماشین به جای نیروی متخصص انسانی در ترجمه متون خلق نشده است، بلکه ابزاری است که سرعت دسترسی به ترجمه متون را تا حد زیادی افزایش داده و میزان بهره‌وری از زمان، فکر و انرژی مترجم‌ها را به شدت افزایش می‌دهد. از دیگر دیدگاه‌های منفی نسبت به این فناوری، مقایسه نتیجه فرایند ترجمه ماشینی با ترجمه انسانی است که بی‌تردید، ماشین‌های مترجم می‌توانند با دقت و ظرافت یک انسان و با توجه به این مطلب که فاقد شعور و درک انسانی هستند، به خوبی عمل کنند.

از آنجا که مسئله‌ی روشها و پارامترهای ارزیابی نتایج ترجمه، همواره در زمینه ترجمه ماشینی مورد توجه بوده‌است، در این مطالعه سعی بر آن بود پس از بیان روش‌های قبلی، روش آماری ترجمه ماشینی تجزیه و تحلیل گردد. از این رو که در گذشته برای ارزیابی کیفیت متن‌های ترجمه شده از داورهای انسانی استفاده می شده است در این تحقیق می‌خواهیم با استخراج الگوهای انسجامی بین یک متن و ترجمه آن، توانایی ماشین را در ارزیابی متن‌های ترجمه شده بررسی کنیم.

در نهایت با توجه به اینکه عقل و اندیشه اساس کار زبان انسان را تشکیل می‌دهد و ماشین فاقد آن است، ریشه مشکلات امروزی ماشین‌های ترجمه را باید در سه مورد جستجو کرد:

1- برنامه‌های نادرستی که به ماشین داده می‌نشود

برای اصلاح برنامه‌ها لازم است در گروه برنامه‌نویسان، زبان‌شناسی مجرب نیز گنجانده شود تا با تحلیل ساختارهای مختلف زبان، الگوهای مناسب را برای تهیه برنامه در اختیار برنامه‌نویسان قرار دهد.

2- ضعف موجود در تئوری زبانی

طبیعی است که نمی‌توان زبان را جدای از معنا بررسی کرد. حتی وقتی درباره آواهای زبانی صحبت می‌کنیم، به معنا نیازمندیم. بنابراین چنانچه تئوری زبانی به گونه‌ای اصلاح شود که معنا را نیز لحاظ کند می‌توان بخش دیگری از مشکلات ماشین‌های ترجمه را حل کرد. البته انجام چنین کاری بسیار دشوار، اما شدنی است.

3- فقدان قدرت اندیشه و تفکر

اگر بتوانیم به شکلی فقدان چنین قدرتی را - یعنی همان نعمتی را که خدا به ما عطا کرده - در اختیار ماشین قرار دهیم، مشکلات ماشین‌های ترجمه را می‌توان رفع نمود. بنابراین باید انتظار داشته باشیم که با توجه به دانش کنونی بشر، در صورتی که مشکلات دسته اول و دوم رفع گردند، مشکلات دسته سوم نیز برطرف می‌شود.

فصل اول

کلیات

1-1- مقدمه

از دیر باز بشر همیشه در آرزوی اختراع دستگاهی بوده است که بتواند متون را از زبان مبداء به زبان مقصد با سرعت و دقت فراوان ترجمه کند. تلاش برای رسیدن به این آرزو بیش از پنج دهه است که آغاز شده است و محققان دستاوردهایی در زمینه طراحی سیستم‌ها و ساخت ماشین‌های مترجم بدست آورده‌اند. ما در این پژوهش سعی داریم اشکالات اساسی این نوع ترجمه را ذکر نماییم. به نظر می‌رسد که با توجه به گسترش روز افزون استفاده از ماشین‌های مترجم در سازمان‌های داخلی و بین‌المللی، آشنایی با این پدیده قرن ناگزیر باشد.

1-2- اهمیت و ضرورت تحقیق

در طول چند دهه اخیر و همزمان با گسترش و پیشرفت زبان‌شناسی محاسباتی، در بسیاری از کشورهای پیشرفته، تلاش‌های همه‌جانبه و پیگیر در جهت ترجمه متون به کمک کامپیوتر انجام گرفته است. با توجه به تنگناها و مسائل خاص مطرح در زمینه ترجمه ماشینی و علی‌رغم کاستی‌ها، حاصل این فعالیت‌ها درخور توجه بوده است. توسعه ارتباطات از طریق اینترنت نیز که بصورت یک بستر جهانی برای مبادله افکار و اطلاعات درآمده است خود مستلزم توسعه ماشین‌های ترجمه می‌باشد تا با سرعت بتوان اطلاعات را از یک زبان به زبان دیگر برگرداند.

رشد جالب توجه شبکه‌های اجتماعی، مانند فیس بوک یا پیام‌رسان‌های فوری همانند اسکایپ¹، گوگل تاک² و ام اس ان مسنجر³، در سال‌های اخیر، مورد استفاده‌ی دیگری برای نرم‌افزارهای ترجمه ماشینی شده است. چراکه به این وسیله کاربران به زبان‌های مختلف می‌توانند با یکدیگر ارتباط برقرار کنند. نرم‌افزارهای ترجمه ماشینی همچنین برای بسیاری از موبایل‌ها، کامپیوترهای جیبی و... عرضه شده است. بخاطر قابلیت حمل آنها، اینگونه وسایل بعنوان ابزارهای ترجمه برای موبایل مشخص شده‌اند که شبکه تجارتي با استفاده از موبایل را بین شرکایی با زبان‌های مختلف ممکن ساخته است. همچنین این نرم‌افزارها نیاز به مترجم انسان بعنوان واسطی در مکالمات دوطرف را از بین می‌برند. (61)

¹ Skype

² Google Talk

³ MSN Messenger

1-3- اهداف مشخص تحقیق

در این پژوهش تلاش بر این است با مطالعه روش‌ها و مدل‌های ترجمه آماری ماشینی الگوهایی پیشنهاد شود که پس از استفاده از این الگوها علاوه بر افزایش سرعت و دقت، متن ترجمه شده متنی منسجم با معنا و مفهومی هرچه نزدیک‌تر به اصل متن در زبان مبدأ باشد.

1-4- پرسش‌های تحقیق

ترجمه‌های ماشینی از چه میزان کیفیتی برخوردارند؟

آیا روش‌ها و مدل‌های موجود قادرند ترجمه‌ی متن را ارزیابی کنند؟

آیا متن ترجمه شده دارای انسجام می‌باشد؟

کدام روش‌ها برای بهبود کیفیت ترجمه مناسب‌ترند؟

1-5- تعاریف واژه‌ها و اصطلاحات

1-5-1- کلمه:

کلمه در یک متن کوچکترین واحد معنادار می‌باشد. بطور معمول بین کلمات در یک متن یک فضای خالی وجود دارد که به آن وسیله می‌توان کلمات را از هم تشخیص داد. البته اگر بخواهیم ترجمه را از روی صوت انجام دهیم، نمی‌توانیم چنین فضای خالی را بین کلمات ایجاد یا بین آنها مکث کنیم و شنونده به صورت پیوسته دریافت می‌کند. به ویژه اگر اصوات به صورتی باشد که شنونده با آن آشنایی نداشته باشد.

1-5-2- جمله:

جمله مجموعه‌ای به هم پیوسته از کلمات است که پیامی را بیان می‌کند. جمله دارای اجزاء و ارکان است. ارکان جمله آن قسمت‌هایی هستند که اگر حذف شوند، ساختمان جمله از هم می‌پاشد. اجزای جمله آن‌ها هستند که حذفشان به ساختمان جمله آسیبی نمی‌رساند. جمله گونه‌های مختلفی دارد که بحث راجع به آن خارج از موضوع این پژوهش است.

1-5-3- عبارت:

گروه یا عبارت مجموعه‌ای از یک یا چند واژه است که در نحو جمله همچون یگانی مستقل عمل می‌کند. هر گروه هسته‌ای دارد که می‌تواند یک یا چند وابسته داشته باشد. بنابراین، وجود هسته در هر گروه الزامیست اما وجود وابسته الزامی نیست. گروه‌ها را بر اساس رده واژگانی هسته‌شان نام می‌گذارند:

- گروه اسمی: هسته این گروه از رده‌ی واژگانی «اسم» است.
- گروه فعلی: هسته این گروه از رده‌ی واژگانی «فعل» است.
- گروه صفتی: هسته این گروه از رده‌ی واژگانی «صفت» است.
- گروه قیدی: هسته این گروه از رده‌ی واژگانی «قید» است.
- گروه حرف اضافه‌ای: هسته این گروه از رده‌ی واژگانی «حرف اضافه» است.

1-5-4- اصطلاح:

اصطلاح‌ها گروه‌هایی از واژه‌ها هستند که در مجموع دارای درک و معنایی ویژه می‌باشند. مفهوم اصطلاح‌ها را نمی‌توان از معنا کردن تک‌تک لغات تشکیل‌دهنده‌ی آنها حدس زد. اصطلاح‌ها کلمه‌ها، عبارت‌ها و بیان‌هایی هستند که به شکلی رایج در گفتگوهای روزمره در زبان‌های مختلف استفاده می‌شوند. اغلب استعاری بوده و به زبان، تنوع و روح می‌دهند. اصطلاح‌ها مربوط به بخش استعاری زبان بوده و با فرهنگ آن زبان عجین هستند. فعل‌های عبارتی نیز از دید معنایی و فهم به نحوی یک اصطلاح محسوب می‌شوند.

اصطلاح‌ها در برخی واژه‌نامه‌ها به صورت برجسته و پر رنگ و در زیر واژه مورد نظر (کلید واژه) قرار می‌گیرند. واژه‌نامه‌های سری لانگمن از این دسته هستند. اصطلاح‌های مترادف به اصطلاح‌هایی می‌گویند که تقریباً هم‌معنای یکدیگر بوده و استفاده از آنها در یک حوزه کاربرد دارد.

به عبارتی دیگر؛ اصطلاح‌ها عبارت‌های ویژه‌ای هستند که تقریباً به وسیله بیشتر افراد یک ملت یا فرهنگ مشخص، شناخته شده، پذیرفته شده و مرسوم هستند مثلاً در فارسی اصطلاح دست شما درد نکند برای همه شناخته شده و مورد پذیرش است.

گاهی معنای اصطلاحات به آسانی قابل حدس هستند و گاهی هم معنای آنها را نمی‌توان به سادگی حدس زد و تشخیص مفهوم و ترجمه آنها با استفاده از کلمات تشکیل دهنده آنها ممکن نیست.

این جمله قابل حدس است (وقت رو تلف نکن). Don't kill the time.

این جمله را نمی‌توان به سادگی حدس زد (تصادف من رو شوکه کرد). The accident curled my hair.

(62)

1-5-5- ترجمه:

صنعت و هنر ترجمه یا برگردان، کار درک و تفسیر موضوعات، معانی، و مفاهیم پدید آمده در یک زبان (زبان مبدأ)، و سپس انتقال، معادل‌یابی و بازسازی آنها در زبانی دیگر (زبان مقصد) را بر عهده می‌گیرد.

کت فورد ترجمه را چنین تعریف کرده‌است:

جایگزینی مواد متنی یک زبان با مواد متنی برابر، در زبان دیگر.

ترجمه به بیان رومن یا کوبسن:

ترجمه یا واقعیت تازه‌ی ادبی صرفاً رابطه‌ی بینامتنی با شعر یا اثر قبلی دارد.

عوامل بسیاری در تبلور معنایی یک متن دخالت دارند و در درک مطلب یک متن، یعنی شناسایی دقیق این عوامل از جانب مخاطب. گاه ممکن است معنای گوینده و معنای مخاطب با هم تفاوت داشته باشد. یعنی آنچه مورد نظر گوینده است با آنچه مخاطب برداشت می‌کند، فرق داشته باشد و یا مخاطبان معانی متفاوتی را از یک متن برداشت کنند. این امر بیان کننده این مطلب است که عوامل بسیاری در داد و ستدهای زبانی حاکم‌اند که متون و ابزارهای نظام زبان (صداها، دستورها و واژگان) تنها یکی از آنها هستند.

انواع ترجمه با توجه به روشی که مترجم انتخاب می‌کند عبارتند از:

1- ترجمه کلمه به کلمه: ترجمه متن به صورت کلمه به کلمه و بدون توجه به ترکیب کلمات و قرار دادن آنها در

جایگاه مناسب با توجه به زبان مقصد.

2- ترجمه معنوی (بهترین نوع ترجمه است): گاهی ترجمه علاوه بر اینکه مطابق با متن می‌باشد، با توجه به زبان مقصد و نحوه ترکیب کلمات در آن زبان انجام می‌شود.

3- ترجمه آزاد (مفهومی): این نوع ترجمه به گونه‌ای است که مترجم، خود را در بند و اسیر عبارات نمی‌سازد و به دلخواه خود و متناسب با وضعیت دست به تغییراتی می‌زند.

4- ترجمه داستانی یا روایتی: گاهی ممکن است مترجم، مفهوم و ریشه متن را گرفته و آن را به شکل داستان بیان کند.

1-5-6- پیشینه ترجمه در ایران:

پیشینه ترجمه در ایران به روزگار هخامنشیان می‌رسد (۵۵۹-۳۲۱ ق.م). گستردگی این شاهنشاهی و ارتباطات گسترده آن با سرزمین‌های دیگر ترجمه را امری ضروری ساخته بود. کتیبه‌های هخامنشیان معمولاً به سه زبان فارسی باستان، اکدی، و ایلامی نوشته می‌شده است.

در مراکز علمی عصر ساسانیان، مانند سلوکیه، تیسفون، و گندی‌شاپور آثار فراوانی از زبان‌های گوناگون متداول در آن روزگار همچون سانسکریت، یونانی، لاتینی، و سریانی به پهلوی ترجمه می‌شد. نخستین ترجمه دوران ساسانیان را مربوط به عصر خسرو اول انوشیروان می‌دانند. از نامدارترین آثاری که در این دوره ترجمه شد، باید به ترجمه پنجه‌تنتره یا کلیله و دمنه از سانسکریت به فارسی پهلوی اشاره کرد.

همچنین، در آغاز تمدن اسلامی ترجمه‌های متعددی از کتب فلسفی و علمی یونانی به پهلوی موجود بود.

کار ترجمه در ایران با کتابهای درسی دارالفنون رونق گرفت و بر اثر احیاء صنعت چاپ و اعزام محصلین به خارج گسترش یافت. از آنجا که هدف تأسیس این مؤسسه تربیتی در زمان عباس میرزا و به پیشنهاد روشنفکران و آزادی‌خواهان، اخذ تمدن و علوم و معارف مغرب زمین عنوان شده بود، لذا به استخدام معلمین خارجی نیز اقدام گردید. همچنین کتب درسی خارجی در زمینه علوم، ادبیات، تاریخ، پزشکی و غیره با کمک همین معلمین به فارسی برگردانده شد. پس از چندی به توصیه عباس میرزا به ترجمه آثار ادبی نیز پرداختند. از آنجا که محصلین بیشتر به فرانسه اعزام می‌شدند زبان فرانسوی در کشور رایج و با گستردگی آن نه تنها آثار نویسندگان فرانسوی بلکه آثار ادبی دیگر زبانهای اروپایی مثل آلمانی، روسی، لهستانی، مجاری، اسپانیایی، پرتغالی نیز به فارسی برگردانده شد.

نخستین نمایش‌نامه‌ای که از فرانسه به فارسی ترجمه شد، نمایشنامه «مردم‌گریز»، اثر مولیر است. دو کتاب تاریخی «پتر کبیر» و «شارل دوازدهم» اولین سفارش ترجمه‌ای بود که عباس میرزا به میرزا رضا مهندس داد. «انحطاط و سقوط امپراطوری روم» سفارش دیگر او به همین مترجم بود. نهضت ترجمه و نشر کتاب پس از مرگ عباس میرزا و فتحعلی شاه، در دوران محمد شاه قاجار به وقفه افتاد، ولی در عهد ناصرالدین شاه و مظفرالدین شاه به اوج خود رسید. محمد طاهر میرزا و محمدحسن خان اعتماد السلطنه پرکارترین مترجمان این دوره بودند. از مشهورترین ترجمه‌های محمد طاهر میرزا «سه تفنگدار»، «کنت مونت کریستو»، «لوئی چهاردهم و عصر و قرنش» را می‌توان نام برد. «طیب اجباری»، «سرگذشت خانم انگلیسی»، «داستان روبنسون سوئیسی» و «شرح احوال کریستف کلمب» نیز از بهترین آثار ترجمه اعتمادالسلطنه به شمار می‌رود. البته اکثر این ترجمه‌ها، بیش از آنکه یک ترجمه باشند، اقتباس، یا تقلید هستند. پس از جنگ جهانی دوم، زبان انگلیسی جای زبان فرانسه را به عنوان زبان دوم گرفت و گسترش روزافزون یافت. اما با وجود اینکه انگلیسی امروزه زبان دوم اکثریت محصلین ایرانی به شمار می‌رود، در قلمرو ادبیات هنوز هم بهترین مترجمان ما اغلب کسانی هستند که از زبان فرانسه ترجمه کرده‌اند.

1-5-7- ترجمه ماشینی:

ترجمه ماشینی زیرشاخه‌ای از زبان‌شناسی محاسباتی بوده که به ترجمه یک متن از زبانی به زبان دیگر، توسط ماشین و بدون دخالت انسان و با استفاده از الگوهای آماری، دانش زبان‌شناسی و الگوریتم‌های کامپیوتری می‌پردازد. ترجمه ماشینی نه تنها دارای جاذبه و کشش علمی بوده، بلکه از دیدگاه اقتصادی و دیگر ضرورت‌های اجتماعی دارای اهمیت بوده و تحقیق بر روی آن را کاملاً توجیه می‌کند. یکی از دلایل استفاده از ترجمه ماشینی، افزایش سرعت و کاهش هزینه است.

1-5-8- انسجام متن:

اما مهمترین ویژگی در یک متن ترجمه شده انسجام ترجمه خروجی است. انسجام در یک متن همه آن روابط معنایی است که میان عناصر سازنده آن متن برقرار می‌شوند و تشکیل سلسله‌ای از کلام را می‌دهد، بنابراین «آنچه متن را یکپارچه و از نوشته‌های پراکنده متمایز می‌سازد، انسجام و پیوستگی آن است. اتصال جمله‌ها یا روابط عوامل

آن با یکدیگر هنگامی پدید می‌آید که تعبیر عاملی در یک بخش از کلام به عاملی دیگر در همان کلام وابسته باشد. اینجاست که رابطه انسجامی شکل می‌گیرد و این دو عامل سبب شکل‌گیری متن می‌شوند.

1-5-9- فرآیند ترجمه:

فرایند ترجمه به شرح زیر است:

- رمزگشایی معنایی متن مبدأ

- کد گذاری دوباره این معنا در زبان مقصد

در پس این فرایند به ظاهر آسان، عملیات شناختی پیچیده‌ای واقع است. به منظور رمزگشایی معنای متن مبدأ، مترجم باید قابلی قابلیت تفسیر و تجزیه تحلیل تمام ویژگی‌های متن را داشته باشد. یک فرایند که احتیاج به دانش عمیقی از دستور زبان، جمله‌شناسی (نحو)، معناشناسی و اصطلاحات از هر دو زبان مبدأ و مقصد دارد به همان اندازه باید دانش مربوط به فرهنگ صحبت کنندگان آن زبان را نیز داشته باشد. از آن جهت، چالشی در ترجمه ماشینی وجود دارد که چگونه یک کامپیوتر را آموزش دهیم که بتواند همانند یک انسان متنی را بفهمد و بتواند یک متن جدید در زبان مقصد بسازد که به نظر می‌رسد توسط انسان نوشته شده‌است. این مسئله ممکن است به روش‌های مختلفی انجام شود

1-5-10- روش‌ها:

برای دست یافتن به ترجمه ماشینی می‌توان از روشی مبتنی بر قوانین زبان‌شناسی استفاده کرد، به این معنی که کلمات از نظر زبان‌شناسی ترجمه خواهند شد. (در واقع متناسب‌ترین کلمات مقصد جایگزین کلمات مبدأ خواهند شد)

این موضوع که موفقیت ترجمه ماشینی پیش از هر چیز نیازمند حل مسئله فهم زبان طبیعی است، اغلب مورد بحث است. به طور عام روش‌های قانون مند، متن را با استفاده از ساختن واسطی سمبلیک - که نهایتاً متن زبان مقصد از آن ایجاد می‌شود - تجزیه می‌کنند.

با توجه به طبیعت نمایش واسطه، یک روش به عنوان ترجمه ماشینی بین زبانی یا ترجمه ماشینی مبتنی بر انتقال معرفی می‌شود. این روش‌ها نیاز به واژگان وسیعی با اطلاعات تک‌واژشناسی، دستور زبان و معناشناسی دارند.