



# داده‌های بنیادین استراتژیک زبان‌شناسی، کلید اکمال متقابل تکنولوژی و هوش مصنوعی در حیطه ترجمه ماشینی

علیرضا ولی‌پور\*

استاد گروه زبان روسی، دانشکده زبان‌ها و ادبیات خارجی، دانشگاه تهران،

تهران، ایران

Email: alreva@ut.ac.ir



چکیده

این پرسش که آیا می‌توان ماشینی اختراع کرد که قادر به انجام رفتارهایی حتی هوشمندانه‌تر از انسان باشد، حتی پیش از پیدایش رایانه‌های قدرتمند و داده‌های هوش مصنوعی نیز مطرح بود. پیش‌تر برخی از اندیشمندان زبان‌شناس و فیلسوفان زبان بر این باور بودند که این کار شدنی نیست. تفوق بر سد موانع زبانی یکی از مهم‌ترین اهداف و وظایف توسعه تکنولوژی ارتباطات است. ایده ساخت یک زبان مصنوعی ارتباطی، مانند اسپرانتو و یا سعی بر یک زبان کردن تبادل اطلاعات هر کدام به عملی نتوانسته‌اند به موفقیت نایل شوند و به‌ناچار ایده اتوماسیون پروسه ترجمه و بهره‌مندی از سیستم‌های ترجمه ماشینی همچنان مورد توجه و مراجعه قرار دارد. ترجمه ماشینی از دستاوردهای مهم هوش مصنوعی و زبان‌شناسی رایانه‌ای است. در این علم از برترین نظریه‌های جدید کاربردی زبان‌شناسی و دانش رایانه‌ای استفاده می‌شود.

امروز ما می‌دانیم که ایده بهره‌برداری از هوش مصنوعی، و تسهیل در بهره‌وری در حیطه‌های مختلف علوم ایده‌ای درست و ضروری بود، اما ابزار تحقق آن مکانیک نبود. مکانیک بستر مناسبی برای ایجاد اکمال متقابل زبان‌شناسی، منطق، آمار و ریاضی بر حول مکانیک نبود و چابکی لازم را برای امکان بهره‌برداری از دستاوردهای خود در تحقق ترجمه ماشینی را فراهم نمی‌کرد. از سوی دیگر با ظهور پدیده‌ای به نام دیجیتال، نه تنها رویه‌های منطق صوری به سمت منطق نوین و ریاضیات به ریاضیات جدید تمایل پیدا کرد، بلکه گرایش‌های نوینی در پژوهش‌های زبان‌شناختی متناسب با حضور در بستر دیجیتال شکل گرفت و به سرعت پیشرفت کرد. در این مقاله سعی شده است پس از ارائه مختصری از تاریخچه موضوع و چهارچوب نظری آن؛ به چگونگی دستاوردهای به‌دست‌آمده پرداخته شود و چشم‌انداز آتی آن با توجه به متدهای انتخاب شده، ترسیم گردد. نویسنده مقاله ضمن ارائه تجربیات شخصی خود در این عرصه، بر ایجاد بانک منابع بنیادی متنوع در حیطه زبان‌شناسی نوین تاکید دارد. به زعم وی، موفقیت‌ها و دستاوردهای آینده، منوط به تقویت داده‌های نوین و مورد نیاز زیرساختی در حیطه زبان‌شناسی، به‌عنوان یک علم بین رشته‌ای با تکنولوژی به‌منظور کسب دستاوردهای کاربردی، اجتناب‌ناپذیر است.

شناسه دیجیتال DOI: 10.22059/JFLR.2021.331469.900

## اطلاعات مقاله

تاریخ ارسال: ۱۴۰۰/۰۷/۱۱

تاریخ پذیرش: ۱۴۰۰/۰۸/۱۲

تاریخ انتشار: پاییز ۱۴۰۰

نوع مقاله: علمی پژوهشی

## کلید واژگان:

دستور تطبیقی، داده

پردازشی، آنفورماتیک، ترجمه

ماشینی، هوش مصنوعی،

دستور ساختاری

ولی‌پور، علیرضا. (۱۴۰۰). داده‌های بنیادین استراتژیک زبان‌شناسی، کلید اکمال متقابل تکنولوژی و هوش مصنوعی در حیطه ترجمه ماشینی. پژوهش‌های زبان‌شناختی در زبان‌های خارجی ۱۱ (۳)، ۵۴۱-۵۵۵.

valipour, A. (2021). The Fundamental and Strategic Linguistic Data, the Key for the Reciprocal Complementation of Technology and Artificial Intelligence in Machine Translation. Foreign Language Research Journal, 11 (3), 541-555.

\* علیرضا ولی‌پور، استاد گروه روسی دانشکده زبان‌ها و ادبیات خارجی دانشگاه تهران، فارغ‌التحصیل گروه زبان‌شناسی مقایسه‌ای، تطبیقی و تاریخی دانشگاه دولتی مسکو است. حیطه‌های مورد علاقه و فعالیت: پژوهش در ترجمه متون کهن روسی، مشکلات ترجمه، ترجمه ماشینی و هوش مصنوعی، آواشناسی همزمانی و درزمانی روسی، زبان‌شناسی تاریخی و تطبیقی و ادبیات روسی و زبان‌شناسی اختصاصی زبان روسی و لغت‌نامه‌نویسی.



# The Fundamental and Strategic Linguistic Data, the Key for the Reciprocal Complementation of Technology and Artificial Intelligence in Machine Translation



**Alireza Valipour\***

Professor of the Faculty of Foreign Languages & Literatures, University of Tehran, Tehran, Iran  
Email: alreva@ut.ac.ir

## ABSTRACT

This question about the possibility of inventing a system more capable than humans in performing behaviors has always been on the table, even before the advent of powerful computers and artificial intelligence algorithms. Some linguists and philologists believed this process to be impossible. Triumph over the obstacles of linguistic barriers is one of the most important goals and tasks for developments in communication technologies. The idea of constructing an artificial communicative language, such as Esperanto, or the attempt at unifying the language for exchanging information has proven to be abortive due different reasons; therefore, the notion of automating translation process and using Machine Translation (MT) systems continues to be of consideration and reference. Machine Translation is one of the important achievements of artificial intelligence and computational linguistics. This field of knowledge contains the latest and greatest theories of applied linguistics and computer science.

Today, we confide in the fact that the idea of exploiting artificial intelligence, and facilitating its application in multiple areas of science, has been a necessity; however, the implementations of this achievement were not provided by mechanics. This field was not a proper context for creating the reciprocal complementation of linguistics, logic, statistics, and mathematics to revolve around it and could not provide the required agility to employ its achievements in accomplishing Machine Translation. On the other hand, with the advent of a phenomenon known as digital, not only did the procedures of formal logic turn into mathematical logic and that of the mathematics turned into modern mathematics, but also new trends were established in this regard and developed rapidly in linguistic researches properly relevant to digital contexts. After a brief introduction to the topic and its theoretical framework, this paper tries to determine how these achievements have been accomplished and intends to elaborate upon the future prospects with respect to the highlighted methods. While presenting his personal experiences in this area, the author of this paper emphasizes establishing a diverse institutional resource in modern linguistics. In his view, future achievements and accomplishments are due to developing the latest and fundamentally required data in linguistics as an interdisciplinary field of knowledge in order to gain practical achievements.

DOI: 10.22059/JFLR.2021.331469.900

## ARTICLE INFO

Article history:

Received: October 3, 2021

Accepted: November 3, 2021

Available online:  
Autumn2021

## Keywords:

*Comparative grammar, Data-processing, Informatics, Machine Translation (MT), Relevant technology, Structural grammar*

valipour, A. (2021). The Fundamental and Strategic Linguistic Data, the Key for the Reciprocal Complementation of Technology and Artificial Intelligence in Machine Translation. *Foreign Language Research Journal*, 11 (3), 541-555.

\* Alireza Valipour, professor of Russian department at the Faculty of Foreign Languages and Literatures, University of Tehran, graduated from the Comparative, Comparative and Historical Linguistics Department of Moscow State University. Areas of interest and activity: Research in translation of ancient Russian texts, translation problems, machine translation and artificial intelligence, simultaneous and contemporary Russian phonetics, historical and comparative linguistics and Russian literature, and special linguistics of Russian language and lexicography.

توقع داشتیم آنها را نیز روسها تألیف و برای بهره‌برداری در سیستم برنامه‌های رایانه‌ای رمزگذاری نمایند!

کارفرمای ایرانی تصور می‌کرد، مجری کارهای اضافی می‌تراشد و به دنبال کسب درآمد بیشتر است و ابراز ناخرسندی می‌نمود. اما در همین جو بی‌اعتمادی، مجری ضرورت تحقق توقع دیگری را خواستار شد که جو بی‌اعتمادی و شک را بیش از پیش افزایش داد.

مسئول تامین داده‌های زبان‌شناسی اعلام کرد تا قبل از تدوین انواعی از فرهنگ‌های بسامدی و لغتنامه متنی و دستوری، با توجه به گرایش تخصصی خاصی که در نوع سیستم ترجمه ماشینی مشخص و سفارش شده بود، نخستین موضوعی که در دستور کار می‌بایست قرار گیرد، ساختن لغتنامه متنی فراگیر بود. مجری روس خواستار آن شد که تمامی متنهای مورد علاقه برای ترجمه می‌بایست در اختیار ایشان قرار گیرد، تا لغتنامه متنی با احتساب کلمات موجود در آن شکل گیرد، زیرا اگر این کار انجام نشود دو اشکال اساسی در مرحله بعدی یعنی تدوین لغتنامه بسامدی و دستوری بروز خواهد کرد، یکی آنکه احصاء کلمات و عبارات میسر نمی‌شد و طبیعتاً متعاقب آن ارزش وزنی خروجی‌ها یکنواخت و غیر قابل درجه‌بندی می‌شد و متعاقب این استراتژی غلط، پارسر (تحلیلگر نحوی) در ارتباط با کلمات و اصطلاحات کلیدی در هنگام ترجمه توأم با خطاهای فراوان خواهد شد...

بعد از چند ماه و پیشرفت ۴۰٪ کار و در حالی که چیزی نمانده بود اولین ورژن آن روی مانیتور ظاهر شود همه چیز تمام شد. شرکت مایکروسافت (Microsoft) به شرکت روسی آرسنال (Arsenal) که ما طرف قرار با آن بودیم پیام داد که انتقال تکنولوژی را نباید بر اساس قوانین تحریمی برای شرکت‌های دانش بنیان ایرانی انجام دهد و گرنه حمایت و همکاری آن شرکت که دیگر آن موقع شعبه‌ای از مایکروسافت در روسیه بود تحت الشعاع این موضوع قرار خواهد گرفت. شرکت آرسنال با توجه به بندی که در قرارداد همکاری با طرف ایرانی داشت به

چند وقت پیش یکی از دانشجویان کارشناسی ارشد که من استاد مشاور پایان نامه وی بودم و همکارم در گروه استاد راهنمای وی بود، برای راهنمایی و مشورت نزد من آمده بود اما در حین بحث شروع به گله کرد که آخر نتایج این تز به چه درد ما می‌خورد که ما صد صفحه مطلب درباره دو حرف اضافه روسی و مقایسه آن با فارسی بنویسیم و اصلاً چه کسی از آن بعداً استفاده خواهد کرد! به آن دانشجو توضیح دادم که ذخیره داده‌های علمی ما بسیار ضعیف است، اینکه ما نمی‌دانیم چگونه از آن داده‌ها استفاده کنیم موضوع دیگری است و بیشتر مربوط به آینده است و دغدغه الان ما نیست، آرشیوهای علمی هر کشوری باید مملو باشد از خروجی، نتایج و دستاوردهای متنوع در تمام علوم. این فکر غلط است که ما تولید علمی را فقط برای الان در نظر بگیریم. چیزی که تاریخ مصرف آن فقط برای امروز باشد فانی است و ذخیره علمی هم اگر برای آینده کوتاه باشد بی‌فایده و ابتر خواهد بود. اندیشه‌های رازی، ملاصدراء، کانت، انیشتین ماندگار است چون فقط تاریخ مصرف امروزی نداشته است.

برای تبیین اهمیت دستاوردهای بنیادی در حیطه زبان‌شناسی نوین یک تجربه شخصی را که در آن حضور داشتم توضیح می‌دهم تا خواننده این مقاله به اهمیت حیاتی داده‌های بنیادی بیشتر و بهتر واقف گردد. مدتها قبل من در بخشی از ساخت پروژه سیستم ترجمه ماشینی روسی به فارسی با متخصصین روسی در این حیطه همکاری می‌کردم.

پروژه ترجمه ماشینی روسی به فارسی، یک پروژه ملی بود و توسط بخش خصوصی و با مشارکت و همکاری یک شرکت بزرگ بعد از مذاکرات دو جانبه متعدد شروع به انجام شد. اما این پروژه به چند دلیل پیشرفت کندی داشت و سرانجام محکوم به شکست شد. مهمترین اشکال این بود که ما، یعنی طرف ایرانی پیش‌نیازهای ضروری بخش زبان‌شناسانه آن را که شامل انواعی از فرهنگ‌های بسامدی و لغتنامه متنی و دستوری بود، نداشتیم و

شرکت مایکروسافت اعلام کرد: قرارداد اجراء به صورت know how است و در واقع خرید دستاوردی از تکنولوژی است و نه انتقال آن. مایکروسافت توضیحات شرکت آرسنال را نپذیرفت و لاجرم آرسنال تسلیم و پروژه تعطیل شد...

باید پذیرفت استراتژی علمی ما در مسیر انجام این پروژه بسیار ناشیانه بود. ما وارد معرکه مشترکی شده بودیم که ابتدا و انتهای آن را نمی‌دانستیم. کارهایی را می‌توانستیم به‌خوبی در حیطه تدوین لغتنامه‌های اولیه متنی، بسامدی و دستوری از قبل آماده نماییم که لازم به سفارش به مجری در بازه فشرده زمانی نباشد. باید بگوییم بهتر است جمله مشهور «خواستن - توانستن است» را مشروط به توأم ساختن خواسته با درایت و ایحاد توانمندی قلمداد نمائیم. از سوی دیگر اتخاذ استراتژی مناسب در انجام فعالیت‌های مشترک با طرف‌های خارجی، به‌خصوص در شرایط تحریم بسیار مهم است. در این تجربه ذکر شده اگرچه طرف قرارداد ایرانی با آرسنال بخش خصوصی بود و نه دولت ما، اما بهتر بود که موضوع لغتنامه‌های آن از اول موارد عمومی و غیر تخصصی می‌بود، ما با اعلام علاقه به موضوع تخصصی بسیار خاصی، عملاً گوش‌های کسانی را تیز کردیم که اصلاً به آن می‌توانست نیازی نباشد، یعنی قبل از اینکه وسیله را به دست بیاوریم، نه غیر مستقیم، بلکه مستقیم گفته بودیم این وسیله را برای چه می‌خواهیم! استراتژی غلط.

قصد من از ذکر این موضوع در مقدمه این گزارش این بود که تأکید کنم داده‌ها و منابع علمی ثروت ذخیره‌ای هستند که روزی از آنها استفاده خواهد شد. نباید این توقع را داشت که بلافاصله و به بهترین شکل یک کالا یا خدمت و یا تولید علمی به دست آید و به قول روس‌ها «مسکو هم یک‌دفعه ساخته نشد»، و مهم آن است که رویه سازندگی در حیطه‌های اصلی و لازم هر کشوری همیشه وجود داشته باشد. من فکر می‌کنم اخلاق امروز ما ایرانیان دارای یک نقص بزرگی شده، عجز و کم‌حوصله هستیم، دنبال برنامه‌های دراز مدت نیستیم. با اینکه در فرهنگ ما به صبر و حلم فراوان توصیه شده است، و ضرب‌المثل «دیگران

کاشتند و ما خوردیم، ما بکاریم و دیگران بخورند» را برای هم نقل می‌کنیم، ولی علی‌رغم این آگاهی اقدامات مؤثر در زمینه‌های حیاتی و مورد نیاز و صد البته در حیطه ایجاد داده‌پردازی‌های نوین بنیادی بسیار کم صورت می‌پذیرد و فقدان نقشه علمی کلان در پژوهش‌های زبانی کاملاً محسوس است.

### بحث و بررسی

ترجمه ماشینی و هوش مصنوعی همواره یکی از بحث‌انگیزترین موضوعات در شاخه‌های زبان‌شناسی نوین و علوم کامپیوتر بوده است. این پرسش که آیا می‌توان ماشینی اختراع کرد که قادر به انجام رفتارهای هوشمندانه‌تر از انسان باشد، پرسشی است که قبل از پیدایش کامپیوترهای قدرتمند و طرح الگوریتم‌های پیشرفته مطرح شده بود. بسیاری از فیلسوفان زبان اعتقاد جدی داشتند که این امر شدنی نیست.

اما به مرور تکنولوژی هوش مصنوعی نشان داد که قادر به شکستن خط‌های قرمزی است که برای آن ترسیم شده است. اما این سؤال مطرح می‌شود که پیشرفت دانش هوش مصنوعی تا چه حد می‌تواند ادامه داشته باشد؟ باید گفت که امروزه دیگر کسی پاسخ منفی به این سؤالات نمی‌دهد و همان‌طور که گفته شد با احتیاط بیشتر، شاید پاسخ نمی‌دانیم بهترین پاسخ به این نوع سؤالات باشد. در حال حاضر سیستم‌های ترجمه ماشینی دچار نقصان‌هایی هستند، اما با این وجود کمک به درک محتوای کلی متن می‌نمایند. (میرونکو ۲۰۱۶، ۱۱۹۴)

اگر چه سابقه مطالعه هوش یکی از موضوعات بسیار قدیمی در تاریخ فلسفه علم است و از دیر باز، فلاسفه سعی داشته‌اند درک کنند که فرآیند دیدن، یاد گرفتن، به خاطر آوردن و استدلال کردن چگونه انجام می‌شود. اما با پیدایش کامپیوترها در اوایل دهه ۵۰ میلادی این امکان عملاً فراهم

گشت که توانایی‌های ذهنی را به قوانین تجربی و تئوری تبدیل نمود. این احساس در بسیاری ایجاد شد که «ابر مغزهای الکترونیکی» جدید پتانسیل نامحدودی برای هوش دارند. اما همان‌گونه که تولید وسیله‌ای برای خلق موجودیت‌های هوشمند مصنوعی ادامه داشت، به مرور ثابت شد بهره‌برداری از رویه‌های هوش مصنوعی بسیار دشوارتر از آن چیزی است که در ابتدا تصور می‌شد و ایده‌های جدید، غنی‌تر و جالب‌تر از اول بروز کردند.

این پرسش که آیا می‌توان ماشینی اختراع کرد که قادر به انجام رفتارهایی حتی هوشمندانه‌تر از انسان باشد، حتی پیش از پیدایش رایانه‌های قدرتمند و داده‌های هوش مصنوعی نیز مطرح بود. پیشتر برخی از اندیشمندان زبان‌شناس و فیلسوفان زبان بر این باور بودند که این کار شدنی نیست.

از سوی دیگر به نظر برخی اندیشمندان در حیطه هوش مصنوعی «از نظر عقلی و طبق روابط آماری و ریاضی و زیرمجموعه‌ای و امکان ترکیبی حروف با یکدیگر، در هر زبانی یک ماشین ساده می‌تواند از ترکیب دو به دو، یا دو به چند و ... کلمات و جملاتی بسازد که بامعنی، بی‌معنا و یا مبهم باشد. مسلماً این ماشین اگر به طور دائم کار کند، می‌تواند در تولیدات خود بهترین اشعار شاعران نامدار و یا جملات نویسندگان مشهور را نیز بنویسد و حتی ممکن است اشعاری بهتر از اشعار فعلی هم بسراید» (پوپف ۱۹۹۰، ۲۰۱). گرچه ماشین این کار را انجام می‌دهد، اما ضایعات آن فراوان است، چون هوش به امانت گذاشته در آن کم است، اما اگر این هوش ارتقاء یابد و حدود ترکیبات برای آن دقیق‌تر تعریف شود، ضایعات کمتر خواهد شد. زیرا فقط انسان است که هوش مبتکرانه دارد و هیچ ابررایانه‌ای از چنین موهبتی، برخوردار نیست. اما انسان می‌تواند با ایجاد شرایط و سازوکارهایی، شرایط تحقق مصنوعی هوش را فراهم سازد، و بسیاری از پروسه‌های هوشمند را به رایانه بسپارد. استفاده از نرم‌افزار و سخت‌افزار این امکان را دست‌یافتنی

ساخته و تنها موضوعی که هنوز در هیچ‌یک از حیطه‌های هوشمند و از جمله ترجمه ماشینی به تکامل لازم نرسیده، ولی با شتاب فزاینده در حال رشد است، مسئله چگونگی طراحی و انتقال بهینه هوش انسان به ماشین است.

پژوهش در حیطه ترجمه ماشینی در طی سالهای اخیر با فراز و نشیب‌هایی توأم بوده است، اما به اعتقاد مارچوک، این حیطه همیشه اهمیت خود را به‌واسطه تمایل انسان برای دستیابی به افق‌های نوین حفظ خواهد کرد. (رانکو ۲۰۲۱، ۱۱۷)

سوابق ارزیابی‌های زبان‌شناسان ایرانی نیز از پدیده ترجمه ماشینی منفی و گاهی نیز مبهم و غلط بوده است. در یکی از آثار سه دهه قبل زبان‌شناسی ایران آمده است که «ترجمه ماشینی به مفهوم و مقیاس تجارتي فعلا قابل عرضه نیست، زیرا نه کیفیت آن رضایت‌بخش است و نه از نظر اقتصادی مقرون به صرفه است.» (باطنی ۱۳۵۴، ۱۰۷) در صورتی که در همان زمان نظام‌های ترجمه ماشینی مانند آمپار در روسیه و سیستم‌ها در کانادا به‌طور حرفه‌ای مورد استفاده قرار گرفته بودند و با شتاب رشد مثبت خود را دنبال می‌کردند.

اما عجیب آن که، مدت‌ها بعد در اولین چاپ فرهنگ اصطلاحات زبان‌شناسی دوباره همین ارزیابی مبهم، کلی و غلط از ترجمه ماشینی ارائه شده است!

خواننده، منفی و مبهم بودن را در پاراگراف زیر می‌تواند مشاهده کند. ساغرونیان در فرهنگ اصطلاحات زبان‌شناسی نوشته است، «ماشین ترجمه یا به عبارت دیگر حسابگر الکترونی! ماشینی است که یک رشته علائم را طبق دستوراتی که قبلاً دریافت داشته به یک رشته علائم دیگر تبدیل می‌کند.» پس از این توضیحات مبهم و کلی اضافه کرده که «پیش‌ویرایش و پس‌ویرایش به کیفیت ماشین ترجمه کمک می‌کند» (ساغرونیان ۱۳۶۹، ۷۷). در صورتی

که این موضوع تا حدی در نظام‌های دوره اول، در برخی مدل‌ها ابداع شده بود و به علت عدم کارایی، به‌خصوص موضوع پیش‌ویرایش آن، به‌کلی منسوخ شد. وی به نقل و استناد از باطنی تکرار کرده است «ترجمه ماشینی به مفهوم و مقیاس تجارتي فعلا قابل عرضه نیست، زیرا نه کیفیت آن رضایت‌بخش است و نه از نظر اقتصادی مقرون‌به‌صرفه است!» (ساغرونیان ۱۳۶۹، ۷۶).

علم بین رشته‌ای از مهمترین آنها که تهیه شبکه واژگانی، لغت‌نامه‌های متنی و پیکره‌های زبانی بهره‌مند نبود.

این روند کند و تا حدی مایوس‌کننده در رابطه با هوش مصنوعی، تا دهه چهارم قرن بیستم ادامه داشت، اما پس از دهه چهارم با اختراع رایانه، با شتاب حیرت‌انگیزی شروع به پیشتازی کرد. پروسه‌ای که همچنان با سرعت و لاینقطع به پیش می‌رود.

هر روز که می‌گذرد، سرعت و ظرفیت سخت‌افزارها افزایش می‌یابد و قیمت‌ها نیز کاهش می‌یابد. پیشرفت در علوم رایانه‌ای به‌طور فوق‌العاده‌ای موفقیت‌آمیز بوده است و به نظر می‌رسد، این رشد فزاینده پایانی نداشته باشد. رایانه‌های موازی می‌توانند محاسبات بسیار سنگین و پیچیده‌تری را به‌شرط دادن محتوای درست و استفاده از دستورات محاسباتی برای به دست آوردن خروجی بهتری انجام دهند. دستاوردهای مشترک هوش مصنوعی با علوم دیگر نیز در حال گسترش است و از کشاورزی تا هوا و فضا را متحول ساخته است. زبان و زبان‌شناسی نیز از این جریان مثبت نمی‌بایست جدا بماند، و این دگردیسی و سرمایه‌گذاری و تغییر یا ارتقاء پژوهش‌های زبانی از جایی می‌بایست شروع می‌شد که متضمن سود و پیشرفت در حیطه زبانی باشد.

اما در مجموع باید گفت، امروز متفکرین و متخصصین در حیطه زبانی در پی دادن پاسخ منفی به این سؤالات نیستند و ارزیابی آنها از چشم انداز آینده این تکنولوژی مثبت است.

ویلهلم لایبنیتس (Leibniz) (۱۶۴۶-۱۷۱۶) اولین کسی بود ایزاری مکانیکی برای انجام ترجمه‌های ساده ساخت. اما، فرموله نمودن منطق او در بستر مکانیک چنان ضعیف بود که ماشین تولید مفهوم او نمی‌توانست داده‌های زبانی را پردازش و نتایج جالبی تولید کند (راسل ۱۳۸۳، ۸).

پیشرفت در این زمینه برای بیش از یک قرن متوقف شد، تا آن‌که چارلز بابیج (۱۸۷۱ - ۱۷۹۲) به این اندیشه رسید که جداول لگاریتمی می‌توانند توسط ماشین محاسبه شوند. او برای این کار ایده‌ای را طراحی کرد که موضوع حافظه قابل آدرس‌دهی، برنامه‌های ذخیره شده و پرس‌های شرطی برای اولین بار در آن گنجانده شده بود. اما این ایده به‌علت قرار گرفتن در بستر مکانیک از نظر کاربردی و تکنولوژیک، پیشرفت چندانی نکرد.

گذر از دهه چهارم قرن بیستم، همان‌طور که اشاره شد با اختراع رایانه، شتاب حیرت‌انگیزی نیز در ارتقاء و تکامل سیستم‌های ترجمه ماشینی در دوره‌های مختلف روی داد، که هر دوره نیز با توجه به به‌دست آمدن پیشرفت‌های قبلی؛ دارای ویژگی‌های خاص آن دوره بوده به شرح زیر بوده است. زبان به عنوان کد مطرح شد. و متعاقب آن لغت‌نامه‌هایی مخصوص این کار ساخته شد. مرحله‌ای که به ساختن اولین اولگاریتم و برنامه نرم‌افزاری منتج شد.

علت را هم باید در آنجا جستجو کرد که در علوم کاربردی غنای سوابق پژوهشی برای نیل به اهداف بزرگ اجتناب‌ناپذیر است. مشابه با آن در امور مربوط به زبان‌شناسی رایانه‌ای، موضوعات مربوط به آن که پیش‌درآمد آنها است، می‌بایست فراهم و آماده گردند. مواردی که علم زبان‌شناسی در قامت یک

اگر چه باید اذعان کرد، نظام‌های ترجمه ماشینی ارائه شده در این دوره، کارایی حرفه‌ای نداشتند و فقط جنبه تبلیغی و معرفی کار را به صورت کلیشه‌ای ارائه می‌کردند. یعنی معلوم بود توانایی این دستگاه تا چه حد محدود است، جمله هم طوری می‌بایست داده می‌شد که ماشین در آنالیز آن به مشکل بر نمی‌خورد، جملات باید به شکل لازم به ماشین داده می‌شد و بعد هم ویراستاری می‌شد.

پس از این دوره مشخصات و ویژگی‌های سیستم‌های ترجمه ماشینی تحول قابل قبولی یافتند. زیرا حالا دیگر در اواسط دهه شصت، ضرورت ساختن نظام ترجمه ماشینی که الگاریتم‌ها در آن نقشی اساسی داشته باشند، به تجربه ثابت شده بود و در بستر آماری توأم با بازدهی و موفقیت قابل قبولی شد. ضمن آنکه تحقق شرح و توصیف مبانی دستوری مناسب ماشین، که با دستور سنتی تفاوت داشت و کلمات براساس معیارهای جدیدی تقسیم می‌شدند، در دسترس کار قرار گرفتند. این معیارهای نوین شامل ویژگی‌های به این شرح بودند:

۱- توجه از لغت بیشتر به نحو معطوف شد. ۲- ایده‌هایی برای تقسیم‌بندی‌های معناشناسی ویژه ماشین مطرح شد. ۳- از الگاریتم‌های نحوی و معنایی در نظام ترجمه ماشینی استفاده شد. ۴- روش‌هایی مدنظر قرار می‌گرفت که امکان ترجمه از یک زبان به زبان‌های گوناگون میسر باشد. ۵- استفاده کاربردی، اما محدود، از ترجمه ماشینی در برخی سازمان‌ها شروع شد. ۶- نیاز به علوم دیگری مثل ریاضی و برنامه‌نویسی و گسترش الگاریتم‌ها احساس شد. ۷- لغت‌نامه‌های بسامدی ساخته شد و برنامه‌ریزی براساس آمارهای آن تدارک دیده شد. ۸- موفقیت کامل در ترجمه صنعتی به دست آمد. ۹- با بالا رفتن سرعت رایانه، امکان بهره‌مندی از برنامه‌نویسی‌های هوشمند فراهم می‌شد.

و سرانجام دوره سوم که از اوایل دهه هفتاد میلادی، همراه با شروع پیشرفت در علم انفورماتیک آغاز گشت.

رشد سریع فناوری رایانه‌ای، امکانات بهتری را در اختیار برنامه‌نویسان قرار داد. البته روند کار تا سال ۱۹۷۶ کند بود (مارچوک ۱۹۸۳، ۲۰).

در اینجا باید به این نکته نیز اشاره کنیم که در دوره سوم؛ دستور نویسان دستور نویسان ساخت‌گرا و فرمالیست‌های جهان را به دو دسته می‌توان تقسیم کرد، که بر اساس آن رویه‌های مانند چامسکی مقابل ابداعات به سبک ملچوک قرار می‌گیرد، و بر اساس تجربه و واقعیات عینی می‌توان گفت، آن‌قدر که در عمل، ایده‌های معناشناختی ملچوک به ترجمه ماشینی کمک کرده، ایده دستور همگانی و جهانی چامسکی کاربرد چندانی برای نظام‌های ترجمه ماشینی نداشته است.

انواع لغت‌نامه‌ها در نظام ترجمه ماشینی نقش‌های متفاوت و مهمی را ایفا می‌کنند، و اولویت‌بندی و ارزش-گذاری مراحل انجام کار، در حیظه واژگان و دستور، بدون تدوین لغتنامه بسامدی حیظه مورد نظر، کاری غلط و بی‌فایده خواهد بود. زیرا برای این که سرعت برنامه تجزیه و تحلیل ماشین به علت عملیات جستجو در حجم کلمات فراوان، کاهش نیابد، از شماره‌گذاری برای بن‌واژه‌ها باید استفاده شود. یعنی بین یا مورفی که بیشترین کاربرد را دارد با شماره ۱ و الی آخر شماره-گذاری می‌شود. شماره‌گذاری نیز براساس داده‌های به دست آمده در لغتنامه بسامدی صورت می‌گیرد، بنابراین دقت آن کامل خواهد بود. (ولی پور ۱۱۷۳، ۱۳۸۶)

همان‌طور که دستور زبان مورد نیاز در این نظام با دستور زبان معمولی متفاوت بوده و دارای گستردگی و پیچیدگی‌های فراوانی است. لغت‌نامه‌های نظام ترجمه ماشینی نیز با لغت‌نامه‌های عادی در ابعاد مختلف متفاوت‌اند، این تفاوت هم در استخراج مدخل لغات، پردازش، کدگذاری و هم در کاربرد آن‌ها برای کاربرد غیر

انسانی و به‌منظور به‌کارگیری در نظام ترجمه ماشینی تدوین می‌گردد. لغت‌نامه‌های موجود در نظام ترجمه ماشینی دارای تفاوت‌های اساسی با لغت‌نامه‌های مورد استفاده عام و غیر ماشینی است، اولاً در این لغت‌نامه‌ها، لغات کدگذاری می‌شوند و براساس کد ذخیره شده، بازخوانی می‌شوند و ثانیاً، دارای این ویژگی‌اند که در مراحل مختلف عملکرد نظام ترجمه ماشینی، به ترتیب و تفاوت مورد استفاده قرار می‌گیرند. کدگذاری و تدوین لغت‌نامه‌های ماشینی باید به‌گونه‌ای باشد که به علت مراجعات متعدد نظام به آن، به موضوع برتری ذاتی آن که امکان بهره‌برداری با سرعت بالا است، صدمه‌ای وارد نشود.

از موضوعات مهم بر سر راه ساخته شدن ترجمه ماشینی در سطح واژگان، مشکل چند معنایی بودن واژه‌ها، شناخته شده‌ترین ویژگی زبان‌های طبیعی است (آپرسیان ۱۹۸۷، ۴۷) و این امر زائیده قانونی است که در زبان‌شناسی تحت عنوان اقتصاد کلامی (economy language) اشتهار یافته است. مترجم به هنگام برخورد به واژه‌ای که دارای چند معنا است، با تکیه بر توانمندی زبانی (competence) خود و بافت و موضوع متن، می‌تواند معادل واقعی واژه را تشخیص دهد و برگزیند، اما رایانه فاقد چنین توانمندی‌ای است. بنابراین طراح ماشین ترجمه ناگزیر است مقدماتی فراهم سازد و در الگوریتم به تمهیداتی دست یازد که از طریق آن بتواند خلأ موجود را ترمیم کند. این تمهیدات باید از نخستین مراحل پردازش، یعنی از مرحله واژگان و نحو و ساختار آغاز شود و تا آخرین مرحله، یعنی مرحله تجزیه و تحلیل معنانشناختی متن تداوم یابد. اما مشکل کلمات چند معنا حل نشده بود، افزون بر آن، علی‌رغم تلاش فراوان کدگذاری و ساختن گرامر مناسب که شامل هم صرف و هم نحو بود، در معنانشناسی ماشینی رشد چندانی روی نداده بود و حتی انتخاب معنا در همکاری انسان و ماشین جنبه تخصصی داشت و بهره‌وری عمومی مورد نظر نبود. بنابراین در کنفرانس ماشین ترجمه که در سال ۱۹۷۷ در لوکزامبورگ برگزار شد، به این

نتیجه رسیدند که روش انسانی و ماشینی به‌طور توأم مورد توجه قرار گیرد.

سرانجام از سال ۱۹۷۷، دست‌اندرکاران نظام‌های ترجمه ماشینی رویه محدودیت را برگزیدند، که طبق آن، همان‌طور که شعر و رمان را در مراحل قبلی برای ترجمه ماشینی ثقیل تلقی کرده بودند، موضوع مترجم کلی هم ایده مناسبی در حیطه این نظام تشخیص داده نشد و تحقق آن به آینده نه چندان نزدیک موکول شد. از نظر عقل و منطق بدون تردید با تقویت دستاوردهای هوش مصنوعی، افزایش سرعت و حافظه رایانه و برخی عوامل دیگر، ترجمه ماشینی هم مانند نان ماشینی، یا خدمات جراحی ماشینی و... جایگزین ترجمه انسانی می‌شوند. اگر چه ممکن است اکنون مضحک جلوه کند اما این امکان دست نیافتنی نخواهد بود که تصورات ژول ورن گونه پوپف تحقق یابد و در کنار ترجمه ماشینی، ادبیات ماشینی نیز به وجود آید، ماشین‌هایی که قادر خواهند بود، داستان و شعر تولید کنند.

در روز ۲۶ ماه می ۱۹۷۸ در نیویورک، سازمان ملل متحد، گردهمایی کارشناسان امور مربوط به ترجمه ماشینی بود و نظام‌های کارآمد معرفی و به نمایش گذاشته شدند. کمپانی کانادایی ورد ترانسلیشن (World Translation of Canada) نظام ترجمه ماشینی خود را تحت عنوان سیستم ارائه داد. این نظام ابتدا در دانشگاه جورج تاون در سال ۱۹۷۴ طراحی شد و سپس با جلب مشارکت مؤسسات تحقیقاتی دیگر تکمیل شد. بسیاری از زبان‌شناسان، خوش‌بینانه استفاده از نظریات دستور گشتاری و ساخت‌های نحوی اندیشمندانی مانند چامسکی را بسیار مهم و کارگشا معرفی کردند، زیرا به اعتقاد ایشان «ترجمه واژه به واژه برای طراحی رایانه‌ای که قادر باشد متنی را از یک زبان به زبان دیگر ترجمه کند، غیرعملی و غیرممکن است» (دبیر مقدم ۱۳۸۳، ۱۲۰).



تردید وجود ندارد که پیشرفت‌های علمی، در حد بسیار زیادی، وابسته به ترجمه و انتشار اطلاعات از طریق آن است. در اوایل دهه هفتاد میلادی در مرکز ترجمه روسیه سالیانه حدود ۶ هزار صفحه مطلب ترجمه می‌شد و در اواخر دهه ۷۰ این رقم به ۱۰۰ هزار صفحه رسید که این تقاضا همچنان در حال افزایش است.

در جامعه اقتصادی اروپا هشت زبان رسمی وجود دارد که همه مدارک و اسناد باید به این زبان‌ها ترجمه شوند. طبق آمار ارائه شده، در سال ۱۹۷۸ تعداد خدمت ترجمه و زبان‌شناس در این سازمان ۱۳۰۰ تن بوده‌اند که در یک سال ۶۰۰۰۰۰ صفحه متن ترجمه کرده‌اند، اکنون مترجمان به ۲۵۰۰ تن رسیده‌اند، ولی باز هم کافی نیست و کمبود وجود دارد (مارچوک ۱۹۸۵، ۵).

امروزه روابط بین الملل هر روز در حال گسترش است. نیاز دانشجویان و متخصصین در حیطه‌های مختلف علوم افزایش می‌یابد. (میرونکو ۲۱۶، ۱۱۹۳)

غیر از نیاز امروز جوامع صنعتی برای دستیابی و تبادل اطلاعات، حذف مانع‌زبانی در ارتباط بین انسان‌ها آرزوی همگانی بشریت بوده است. بنابراین انگیزه‌های کافی و قوی و حتی دائمی برای تحقق این هدف وجود داشته و دارد. هیچ فرد و یا دولتی بی‌نیاز از کسب اطلاعات مختلف نیست، امروزه داد و ستد اطلاعات، گسترده‌تر و سودمندتر از گذشته شده است و این روند ادامه خواهد داشت. تلاش برای ساخت نظام‌های ترجمه ماشینی در روسیه و غرب، سابقه‌ای فراتر از نیم قرن دارد، با این حال، روند این تحقیقات و تجربیات دارای فراز و نشیب و تغییر و اتخاذ سازوکارهای مختلف بوده است. از سوی دیگر، علم ترجمه ماشینی که تاریخچه کاربرد عملی و حرفه‌ای آن، مانند موضوع هوش مصنوعی، کمی افزون‌تر از نیم قرن است، محل تلاقی و در هم تنیده شدن علم انفورماتیک و به-خصوص مهندسی نرم‌افزار با زبان‌شناسی است. این

یکپارچگی به‌گونه‌ای بوده که در برخی موارد مرز بین این علوم فرو ریخته است و برخی زبان‌شناسان و یا مهندسان نرم‌افزار به تنهایی و مستقل از یکدیگر سعی در تحقق این هدف داشته‌اند. (ولی پور ۱۳۸۶، ۱۵۷)

تکنولوژی ترجمه ماشینی با قدمت هفتاد ساله مملو و توأم از تجربیات موفق و ناکام بوده است. (پاناسنکف ۲۰۱۹، ۱۰)

با این حال تجربه‌های موفق این کار، حاکی از مشارکت زبان‌شناسان با مهندسان است، ضمن آنکه اطلاع از سازوکارهای طرفین و رعایت تقسیم کار تخصصی، موجب تسهیل و ارتقاء کار در حیطه سیستم‌های هوشمند جهت بهره‌برداری در ترجمه ماشینی خواهد شد.

امروز ما می‌دانیم که ایده بهره‌برداری از هوش مصنوعی، و تسهیل در بهره‌وری در حیطه‌های مختلف علوم ایده‌ای درست و ضروری بود، اما ابزار تحقق آن مکانیک نبود و نمی‌تواند باشد. مکانیک بستر مناسبی برای ایجاد اکمال متقابل زبان‌شناسی، منطق، آمار و ریاضی بر حول مکانیک، نبود و چابکی لازم را برای امکان بهره‌برداری از دستاوردهای خود در تحقق ترجمه ماشینی فراهم نمی‌کرد. از سوی دیگر با ظهور پدیده‌ای به نام دیجیتال، نه تنها رویه‌های منطق‌صوری به سمت منطق نوین و ریاضیات به ریاضیات جدید تمایل پیدا کرد، بلکه گرایش‌های نوینی در پژوهش‌های زبان‌شناختی متناسب با حضور در بستر دیجیتال شکل گرفت و به سرعت پیشرفت کرد. (ولی پور ۱۲۲، ۱۳۸۶)

بعد از آن زبان‌شناسی معاصر به واسطه پیشرفت روزافزون در عرصه ترجمه ماشینی و احتیاج گسترده به این تکنولوژی، ضرورت پرداخت به پژوهش‌های نوین در جمع‌آوری و پردازش داده‌های زبانی، اجتناب‌ناپذیر می‌نماید. پردازش کلمات و تجزیه و تحلیل صرفی واژه‌های

هر زبان ویژگی‌ها و مشکلات مخصوص به خود را دارد. بعضی از این مشکلات در همه زبانها به صورت مشترک وجود دارد و به زبانی خاص اختصاص ندارند؛ مشکلاتی مانند: چند معنایی، تعیین سبک و مشخص کردن حدود آن، چون طبیعی است که برگرداندن جمله‌ای به زبان دیگر لزوماً توأم با تساوی نشانه‌های صرفی و زنجیری نخواهد بود، مفهوم کلمات نیز از زبانی به زبان دیگر، کاملاً همگن نخواهد بود. (بارخودارف (Бархударов)، ۱۹۹۵، ۱۲۲) برای مثال، فرض کنید فهرستی نامتناهی از جمله‌های زبان انگلیسی به همراه توصیف ساختی و نحوی آن جمله‌ها در اختیار یک فرد انگلیسی زبان گذاشته شود. آن فرد با داشتن فرهنگ لغت زبان انگلیسی و با بهره‌جویی از دانش زبانی خود قادر است معنای هر یک از جمله‌های آن فهرست نامتناهی را به دست دهد. او قادر است معانی مختلف یک جمله را تشخیص دهد، بگوید آیا جمله‌ای بی‌معناست یا بامعنا، و جمله‌های هم‌معنا در آن فهرست را نیز بازشناسد. حال وضعیت فوق را با شق زیر مقایسه کنید. فرض کنید دستگاهی ساخته‌ایم که به گونه‌ای مکانیکی قادر است با استفاده از یک فرهنگ لغت، معنی یا معانی هر تکواژ در هر یک از جمله‌های این فهرست نامتناهی را ثبت کند. آشکار است که تعداد معنایی که برای هر تکواژ در فرهنگ قید می‌گردد به مراتب افزون‌تر از معنی آن تکواژ در هر جمله است. اما دستگاه فرضی ما قادر نیست معنی صحیح آن تکواژ را در یک جمله به دست دهد، مگر آنکه معنی مورد نظر از طریق اطلاعات نحوی مربوط به آن تکواژ تعیین گردد. به عنوان مثال:

جملاتی مانند *the fish was bought by the*  
*cook, the fish was bought by the river*. عدم  
 درک ماشین را در چنین موقعیت‌هایی متذکر می‌شود، در صورتی که می‌توان با نگارش لگاریتم مناسب به نظام ترجمه ماشینی، دستور داد کجا کلمه *by* را قید و کجا نشانه مجهول را به کار برد. و یا در نمونه دیگر که در زیر آوردیم

دستگاه قادر است معنی صحیح واژه *seal* را در جمله *seal the letter* برگزیند مشروط بر اینکه اطلاع داشته باشد این واژه در این جمله فعل است. به همین ترتیب، دستگاه با داشتن این اطلاع که واژه *seal* در جمله *The seal is on the letter* اسم است قادر است معنی صحیح این واژه را به دست دهد. اما این دستگاه قادر نیست معنی درست *seal* را در جمله *one of the oil seals in my car is leaking* (که «واشر روغن» می‌باشد) از معانی نامربوط آن نظیر «مهر»، «نقش و تصویر یک مهر» و غیره بازشناسد و این اطلاع دستوری که این واژه در این جمله اسم است نیز کمکی به روشن کردن معنی در اینجا نمی‌کند زیرا کلیه معنایی که برای *seal* در این مثال برشمرديم معانی اسمی این واژه‌اند. آنچه دستگاه ما از انجام دادن آن ناتوان است بهره‌جویی از روابط معنایی بین تکواژهای یک جمله است، از این رو نمی‌تواند معنی یا معانی صحیح یک جمله را تعیین کند. و نیز به خاطر همین ناتوانی دستگاه ما نمی‌تواند جمله‌های بی‌معنا را از جمله‌های معنادار بازشناسد. به عنوان نمونه، این دستگاه نمی‌تواند بین جمله *The wall is covered with*  
*silent paint* و جمله *The wall is covered with fresh*  
*paint* تمیز قائل شود. سرانجام اینکه، این دستگاه تنها جمله‌هایی را که از نظر نحوی عیناً همساخت‌اند و واژه‌های سازنده آن جمله‌ها نیز دقیقاً یکسان‌اند و یا مترادف، به عنوان هم‌معنا تشخیص خواهد داد! (دبیر مقدم، ۱۳۸۳، ص ۱۲۰).  
 بنابر این علاوه بر ترتیب واژه‌ها، توالی واژه‌ها ممکن است به علت تخطی از قواعد نحوی مربوط به گزینش مقوله‌های واژگانی مناسب نیز غیر دستوری شود. برای تایید این موضوع نگاه کنید، چگونه در مثال جمله (a) دستوری، اما جمله (b) غیر دستوری است.

- الف . مزه کیک خوب است.  
 a . The cake is good .

b. - ب. این جمله از نظر معنایی اشتباه است.

The expire is good .

اشکال جمله (ب) در این است که قواعد نحوی زبان انگلیسی ایجاب می‌کنند که پس از **the** و پیش از **is** مقوله یا طبقه خاصی از واژه‌ها باید ظاهر شود؛ این مقوله را در دستور سنتی با برچسب اسم مشخص می‌کنند. در جمله (الف) واژه **cake** اسم است، در نتیجه جمله دستوری است، اما در جمله (ب) واژه **expire** فعل می‌باشد.

واژه **cake**، به یک مقوله تعلق دارد که در برابر آن واژه‌ای مانند **expire**، گروه دیگری از واژه‌ها را تشکیل می‌دهد. برای نامیدن این قبیل مقوله‌ها در طول زمان برچسبهایی ایجاد شده است، و نخستین واژه فوق نمایانگر اسامی، و دومی حاکی از افعال هستند، دانش مربوط به مقوله‌های واژگانی از معیارهای نحوی است که سخنگویان را قادر می‌سازد تا جملات دستوری را تولید و درک کنند، و هم‌زمان با آن، جملات غیر دستوری را مشخص کنند و از آنها اجتناب ورزند. از این رو، توصیف نحو زبان انگلیسی باید شرح مقوله‌های واژگانی را نیز در برگیرد. بنابر این واضح است که به علل کم و بیش مشابه اشکالاتی در ترجمه‌های ماشینی نیز تا مدتها قبل از تکامل آنها بروز خواهد کرد. زیرا نحو و معنا که برخی این دو را حتی یک مقوله می‌دانند، دشوارترین بخشهای هر زبان را تشکیل می‌دهند، که برخی از بزرگان معاصر علم زبان‌شناسی مانند بلومفیلد فرمولیزه کردن دقیق آن را برای کاربردهای عملی غیرممکن تلقی نموده و تلاش در این زمینه را بی‌فایده دانسته‌اند. با این حال اکنون مشهود است با پیشرفت در حیطه ترجمه ماشینی، جایگاه ترجمه انسانی تنزل خواهد یافت. (بایوی ۲۰۲۱، ۷۹)

یادآور می‌شویم، هنگامی که زبان‌شناسان ساخت‌گرا بررسی نحو را آغاز کردند، روالهای کشف که به نظر می‌رسید در مورد آواها و واژه‌ها موثر افتاده اند، دیگر نتیجه‌ای

به دست ندادند. بررسی جمله‌سازی، که به مراتب پیچیده‌تر از آن آواها و واژه‌سازی است، مستلزم اصول و روشهای پژوهشی و تجزیه و تحلیل پیچیده‌تری بود. زبان‌شناسان ساخت‌گرا در درک ما از معنا سهم اندکی داشتند. (جولیا اس فالک ۱۳۷۷، ص ۳۰) به همین علت دستاوردهای ایشان کمکی به ساخت الگوریتم‌های مناسب برای ترجمه ماشینی نکرد. به اعتقاد نویسنده این مقاله، ارزیابی از دستیابی به نتایج فرمولیزه کردن دقیق نحو برای کاربردهای عملی توأم با افراط و تفریط بوده است. تجربه نشان می‌دهد نه پاک کردن صورت مسئله به سبک بلومفیلدیان و هم ساده انگاریهایی از استفاده و تعمیم‌های کلی روش‌هایی مانند چامسکی و دیگر ارائه‌کنندگان دستورهای جدید کمکی به ارتقاء بهبود ترجمه ماشینی نتوانسته نماید. تاریخ کوتاه زبان‌شناسی کاربردی نشان می‌دهد، ساده انگاری از دستاوردهای تحولات انقلابی نه تنها کمکی به پیشرفت این علم نکرده بلکه ارتقاء واقعی آن را به تاخیر انداخته است. این پروسه دقت و حوصله فراوانی را می‌طلبد و ساده لوحانه است اگر تکمیل و تکامل آن را فقط منوط به ابداع نظریات جدید زبان‌شناسی بدانیم. به‌عنوان نمونه رابرت لیز پس از انتشار کتاب ساخت‌های نحوی چامسکی، از امکان بهره‌گیری این نظریه در زمینه «ترجمه ماشینی» (**machine translation**) با اهمیت و خوش‌بینی فراوان سخن گفته بود. وی شیوه‌های موجود آن زمان (یعنی ترجمه واژه به واژه) برای طراحی رایانه‌ای که قادر باشد متنی از یک زبان را به زبان دیگر ترجمه کند غیر ممکن و غیر عملی می‌دانست، و در عوض ساختن سه دستگاه مجزا را که هر یک منطبق با یکی از (صورت تکمیل شده) سه سطح زبانی در نظریه چامسکی است به‌عنوان یک امکان منطقی و محتمل در زمینه ترجمه ماشینی تلقی کرده بود (دبیر مقدم، ۱۳۸۳ ص ۱۲۰). پیشرفت‌های بعدی در زبان‌شناسی نشان داد که مسئله نسبت به آنچه که به نظر می‌رسید، بایست به‌طور قابل ملاحظه‌ای پیچیده‌تر باشد. زبان دارای ابهامات و

ناگفته‌های زیادی است. این بدان معنی است که فهم زبان نیاز به فهم موضوع و زمینه آن دارد و نه تنها یک فهم ساده از ساختار جملات. (راسل ۱۳۸۳، ص ۱۶)

چامسکی در کتاب خود چند اثر را نام برده است که دستور زایشی را ماحصل تلاش‌هایی برای استفاده از رایانه برای اهداف مختلف همچون ترجمه ماشینی دانسته‌اند، ولی او این موضوع را رد می‌کند و ریشه‌های این دستور را در زبان‌شناسی سنتی می‌داند. (دبیر مقدم، ۱۳۸۳ ص ۱۷۹). با این حال چامسکی از ابراز دقیق انگیزه خود برای ارائه این دستور خودداری نموده است. و مستندات وی و پیروانش در توجیه برتری این دستور نسبت به دستورهای دیگر، دلیلی بر کاربردی بودن آن در حیطه ترجمه ماشینی نمی‌تواند باشد. چامسکی مدت‌ها قبل به این باور اعتقاد یافته بود، که مقایسه جمله‌های هسته‌ای در زبان‌های مختلف نشان خواهد داد، به‌رغم تفاوت ظاهری و روبنایی این زبان‌ها، در سطح جمله‌های هسته‌ای تشابه فراوان بین آنها وجود دارد و ابراز امیدواری کرده بود با رمز گشایی همگانی‌ها در ژرف ساخت جملات، زمینه برای ساخت الگوریتم‌های مناسب طبق روش‌های مرسوم شده در دستور گشتاری فراهم آید. ولی پور ۱۳۸۶، ۱۱۹

امروز به دلایل مختلف، که به برخی در این مقاله نیز اشاره می‌شود، معلوم گشته است این تصور عجولانه و تا حد زیادی خوش‌بینانه بوده است و عملاً چنین چیزی تحقق نیافته است. زیرا اشکالات در ترجمه ماشینی معمولاً آن‌چنان‌که در مقدمه نیز اشاره شد، هنگامی بروز می‌کند، که سازوکار ماشینی قادر به تشخیص روابط معنایی بین تکواژهای یک جمله نیست، از این رو نمی‌تواند معنی یا معانی صحیح یک جمله را تعیین کند. ضمن آنکه گشتارهای مختلفی که تاکنون بیانگر تبدیل ژرف‌ساخت به روساخت است، از نظر کمی و کیفی تمام ساخت‌های نحوی را پوشش نمی‌دهند و استثنائات و نکات مبهم

فراوانی هنوز وجود دارد. از سوی دیگر ارزیابی علمی از ترجمه انسانی و ماشینی کار ساده‌ای نیست و امروزه از مسائلی که در کارکرد ترجمه ماشینی مورد توجه است، اتخاذ ساز و کارهایی است که ارزیابی مناسبی از نتایج ترجمه ارائه نماید. ارزیابی سیستم‌های ترجمه ماشینی با دادن یک یا چند جمله و مقایسه آن با ترجمه خوب انسانی، ساده‌ترین روش محسوب می‌شود که قابل انجام توسط افراد غیرمتخصص نیز می‌باشد، اما چنین ارزیابی‌هایی دقیق و زبان‌شناسانه نیست و بیشتر به‌عنوان قرینه و متمم تحقیقات تخصصی مورد استفاده قرار می‌گیرد.

نظام ترجمه ماشینی شامل در حیطه زبان‌شناسی شامل تحقیق و بررسی درباره مقولات واژگان، صرف - نحو، و معناشناسی است. مؤلفه‌های سه‌گانه اشاره شده در تناظر با سه سطح که علم زبان‌شناسی نوین برای هر زبان طبیعی و زنده قائل است، قرار دارند. نحوه بررسی و ارائه دستاوردهای پژوهش در این ابعاد با روش‌های سنتی و آموزشی مرسوم متفاوت است. می‌توان چنین نتیجه گرفت، که در ساخت ترجمه ماشینی، محدودیتی در استفاده از روش‌ها و نظریه‌های زبانی وجود ندارد. برای ساخت لغت‌نامه‌های متنی و دستوری و تحلیلگر نحوی از دستاوردهای دستور زبان تطبیقی (Comparative grammar)، مقابله‌ای (Contrastive grammar)، احتمالی (Expectance grammar)، محوری (Pivot grammar)، ساختاری (Structural grammar)، ساخت‌گروهی (Phrase structure grammar)، گشتاری (Transformational - Generative grammar) و هر روش و نظریه جدیدی که مناسب ساز و کار آن باشد، تاکنون در نظام‌های مختلف ترجمه ماشینی استفاده شده است. برای نمونه، نحوه مشخص کردن بن کلمه، و یا ریشه فعل در نظام ترجمه ماشینی با تعاریف سنتی و آموزشی رایج در رابطه با زبان روسی متفاوت است، و اگرچه در نظام ترجمه ماشینی موضوع کلمات

چند معنا را می‌توان با لغت‌نامه‌های متنی حل کرد، اما همیشه این روش نتیجه‌بخش نیست. مواردی وجود دارد که باید رویه دیگری برگزید. فعالیت و پژوهش‌های زبان-شناسانه و داده‌پردازی زبانی در حیطه زبان‌شناسی رایانه‌ای هر روز افزایش می‌یابد و تردیدی وجود ندارد که دستاورد این تلاش‌ها در زمینه تدوین لغتنامه و دستورهای مورد نیاز به بهبود ارتقاء کیفیت نظام‌های ترجمه ماشینی منجر خواهد شد. (کوزمین ۲۰۱۹، ۱۱۳)

همه این مشخصه‌ها دال بر آن است که میدان پژوهش و تحقیق در این حیطه گسترده است و نیاز به توجه در ایجاد زیر بنای لازم در ابعاد مختلف آن، از ضروریات زبان‌شناسی معاصر در ایران و جهان است.

### تیجه گیری

نظام ترجمه ماشینی فعالیتی چند بعدی است، که در چند حیطه مبتنی بر ابتکار عمل و قدرت حل مسئله است. در این مهارت از برخی دانش‌های بین رشته‌ای نیز استفاده فراوانی می‌شود. ترفندهای هوش مصنوعی با به‌کارگیری اصول ریاضی، آمار و منطق، قالب مناسبی از زبان را که مناسب برنامه‌نویسی و انتقال به ماشین است، فراهم می‌آورد. برای پیشرفت در حیطه ترجمه ماشینی نیازمند احاطه در حیطه‌های داده‌پردازی، توصیف و ساختن مدل‌های مناسب زبان برای کاربرد هوشمند، ساخت و استفاده از نرم‌افزارهای مربوط به این کار و همچنین به کار گرفتن تکنولوژی مرتبط با هوش مصنوعی و برنامه‌نویسی. در این مقاله با لحاظ قرار دادن تفکیک این سه زمینه، برخی ابعاد زبان‌شناختی نظام ترجمه ماشینی، براساس محور قرار دادن روش‌های تجربه شده تبیین شد.

به‌واسطه پیشرفت روزافزون در عرصه ترجمه ماشینی و احتیاج گسترده به این تکنولوژی، ضرورت پرداخت به پژوهش‌های نوین در جمع‌آوری و پردازش داده‌های زبانی، اجتناب‌ناپذیر

می‌نماید. اگرچه سابقه ساخت سیستم‌های ترجمه ماشینی امروزه به بیش از پنجاه سال می‌رسد، ولی علی‌رغم تلاش‌های فراوان در سطح تطابق‌های واژگانی و صرفی، هنوز پیشرفت چندانی در حیطه‌های نحوی و معنا شناسی دستوری (و نه واژگانی) به دست نیامده است. آنچه مسلم است توصیفات سنتی دستور سنتی در سطح نحو، ابهامات تطبیقی را در هنگام ترجمه مرتفع نمی‌کند. طبق نظریاتی که مورد تأیید تأویل‌گرایانی، مانند چامسکی بود، ژرف‌ساخت تمام زبانها نظیر هم هستند، و زبانها فقط در گزینش تأویل با هم اختلاف دارند. این موضوع هنوز به علت ابهاماتی که در عمل بروز می‌کند، به‌صورت یک قانون قطعی تجویز نمی‌گردد.

اگر این موضوع کاملاً مصداق می‌داشت، متضمن بازتاب‌های خوب و کاربردی در کار ترجمه ماشینی می‌شد، یعنی در آن صورت الگوریتم نحوی سازوکارهای ماشینی صرفاً باید بر این موضوع متمرکز می‌شد تا پس از تعیین تأویلهای زبان مبدا و مقصد و حذف روساخت جملات ژرف‌ساخت همگانی به دست آید و سرانجام تأویلهای زبان مقصد را برای به دست آوردن ترجمه، بر ژرف ساخت به‌دست‌آمده اعمال کرد. اگر چه در دوره‌های مختلف نتایج فرمولیزه کردن دقیق نحو برای کاربردهای عملی توأم با افراط و تفریط بوده است. اما تجربه نشان می‌دهد نه پاک کردن صورت مسئله به سبک بلموفیلدیان و هم ساده‌انگارهایی از استفاده و تعمیم‌های کلی روش‌هایی مانند چامسکی و دیگر ارائه‌کنندگان دستورهای جدید کمکی به ارتقاء بهبود ترجمه ماشینی نتوانسته نماید.

بنابر این برای ارتقاء سازو کارهای ماشینی، چاره‌ای باقی نمی‌ماند، که به موضوع کشف و استخراج روابط نحوی در ابعاد گسترده‌تر و با جزئیات مورد نیاز در نحو و معناشناسی پرداخته شود. موفقیت ترجمه ماشینی در آینده منوط به تلاش در شناخت و استخراج روابط معنایی و ساختارهای نحوی و دسته‌بندی و فرمولیزه کردن آنها برای کاربرد در این حیطه است. تجربه ما در این پژوهش نیز تأیید این موضوع بود، که اگر سازوکار ماشینی

۹- راسل، جی. اس. (۱۳۸۳). هوش مصنوعی، رهیافتی مدرن. مترجم رامین رهنمون و آناهیتا هماوندی. چاپ سوم. تهران: انتشارات ناقوس.

۱۰- ساغرونیان، سید جلیل. (۱۳۶۹). فرهنگ اصطلاحات زبان‌شناسی، مشهد: انتشارات نما.

## Resources

1. Falk, Julia S. (1998). Linguistics and language. Trans. Gholam Ali Zadeh, k. Ed. 5th . Astan Print.

2. Dabir Moghaddam, M. (۲۰۰۴). Theoretical linguistics. Tehran: Samt.

3. Russell. Robert. (2004). Artificial intelligence. Tr. Rahnamoun, R. Tehran: Naghoos Press.

4. Valipour, A. An examination of the linguistic dimension of machine translation and a brief introduction of relevant sciences. Foreign Languages Research, No 37, Spring 2007.

5. Valipour, A. A comparison and review of some Russian and Western machine translation systems. Foreign Languages Research, No. 42, Winter 2007.

6. Bateni, M. (1976). Modern linguistic issues, Tehran: Agah.

7. Dabir Moghaddam, M. (2004). Theoretical linguistics, Tehran: Samt.

8. Deitel, Harvey M. (2006). XML How to Program. Tr. Pashaie, B. Ed. 11th. Tehran: Jahanno.

9. Russell, J. S. (2004). Artificial intelligence, modern approach. Tr. Rahnamoun, R and A, Hamavandi. Ed. 3rd. Tehran: Naghoos Press.

10. Sagharvanian, S. J. (1990). The dictionary of linguistic terms, Mashhad: Nama.

هرچه بیشتر امکان فراهم شده از الگوریتمهای نحوی برایش فراهم شده باشد، ترجمه‌های موفق‌تری ارائه خواهد نمود. مصداق عملی این ادعا در عملکرد و ارزیابی‌های انجام شده از سازوکار نظام‌های ترجمه ماشینی به تجربه ثابت شده است. دستیابی نسبی به دستاوردهای مورد نظر، باعث تقویت بنیادین در داده‌های زبان‌شناسی نوین و موجبات استفاده کاربردی آن در تکنولوژی را فراهم خواهد کرد.

## منابع

۱- جولیا اس فالک. (۱۳۷۷). زبان شناسی و زبان. مترجم خسرو غلامعلی زاده. چاپ پنجم. مشهد: موسسه چاپ و انتشارات آستان قدس رضوی.

۲- دبیر مقدم، محمد. (۱۳۸۳). زبان شناسی نظری. تهران: نشر سمت.

۳- راسل. رابرت. (۱۳۸۳). هوش مصنوعی. مترجم رامین رهنمون. تهران: نشر ناقوس.

۴- ولی‌پور، علیرضا. بررسی بعد زبان‌شناختی ترجمه ماشینی و معرفی اجمالی علوم مرتبط با آن. پژوهش زبانهای خارجی، شماره ۱۳۸۶، بهار، ۳۷

۵- ولی‌پور، علیرضا. مقایسه و بررسی برخی نظامهای ترجمه ماشینی روسی و غربی. پژوهش زبانهای خارجی، شماره ۴۲، زمستان ۱۳۸۶

۶- باطنی، محمدرضا. (۱۳۵۴). مسائل زبان‌شناسی نوین، تهران: نشر آگاه.

۷- دبیر مقدم، محمد. (۱۳۸۳). زبان‌شناسی نظری، تهران: انتشارات سمت.

۸- دیتل، هاروی. (۱۳۸۵). راهنمای جامع برنامه‌نویسان. مترجم بهرام پشایی. چاپ یازدهم. تهران: انتشارات جهان نو.

11. Apresyan, Yu.D. (1987). *Linguistic support in the third generation automatic translation system. USSR Academy of Sciences*, preprint of the scientific council on “the problem of Cybernetics”. M.
12. Yngve, V. (1965). The value of machine translation research. – ITY, M.
13. Kuzmin O. I (2019) Collegium Linguisticum -Prospects For The Development Of Automated Translation Systems Using The Specifics Of Pre- And Post-Editing (By The Material Of The English Language), 133-142
14. Marchuk, Yu.N. (1983). *Machine translation problems*. M.: Hayka.
15. Melchuk., I.A. (1989). *To the construction of a working model of the language “Sense-Text”*. M.
16. Mironenko A.B, Omelchenko P.A. (2016) Problems Of Machine Translation Seriya 2, 1193-1195
17. Panasenkov N.A. (2019) Evolution Of Machine Translation: From Babbage’s Analytical Engine To Neural Networks : 9-13
18. Rarenko Maria B (2021) Machine Translation as a challenge. 1. Inion of The Russian Academy of Sciences (Ras), Moscow, Russia: 117-126
19. Popov, E. V. (1990). E.V. (1990). *Artificial intelligence book 1 radio and communication*. M.
20. Baiwei Liu ( 2021) The Future Of Machine Translation And Human Translation In The Belt And Road Initiative Vestnik Moskovskogo Universiteta. Seriya 2, 79-88
21. Barkhudarov T.I. New lexicography. — M. : VTsP, 1995.—183 c.