

## مقایسه و بررسی برخی نظام‌های ترجمه ماشینی روسی و غربی<sup>۱</sup>

علیرضا ولی‌پور\*

دانشیار دانشکده زبان‌های خارجی دانشگاه تهران، ایران

(تاریخ دریافت: ۸۶/۶/۲۷، تاریخ تصویب: ۸۶/۸/۶)

### چکیده

نظام‌های ترجمه ماشینی برتر دنیا قادرند کتابی هزار صفحه‌ای را در کمتر از ۲ دقیقه ترجمه کنند، کاری که مترجم زبده و حرفه‌ای حداقل یک سال وقت روی آن صرف می‌کند. منتقدان، به کیفیت ترجمه این نظام‌ها ایراداتی دارند، با این حال روند ترجمه ماشینی به سرعت رو به رشد است و باید به یاد داشت، ترجمه در نوع سنتی (انسانی) نیز رشد خود را در طی قرن‌ها به دست آورده است و اکنون نیز این‌گونه ترجمه‌ها بدون نقص نیستند. ترجمه‌های قدیمی اکثراً بسیار بد ترجمه شده‌اند و نقد و نظارت و کنترل نیز در کار نبود. فردی مثل ابوعلی سینا ۴۰ بار کتاب مابعدالطبیعه ارسطو را خوانده و نفهمیده بود و این علتش بدی ترجمه بوده است. در این مقاله، عملکرد نظام‌های ترجمه ماشینی برتر روسی و غربی با هم مقایسه شده‌اند، تا علل برتری و نقاط ضعف آن‌ها نسبت به یکدیگر مشخص شوند. در مواردی این ترجمه‌ها با ترجمه‌های انسانی مقایسه شده است. همچنین تلاش شده پس از شناسایی علل اساسی غیر کارآمد بودن نظام‌های ترجمه ماشینی، راهکارهای ارتقاء بهبود کیفیت ترجمه به طور مسند ارائه شود.

**واژه‌های کلیدی:** ترجمه ماشینی، دستور گشتاری، روابط معنایی، الگاریتم، معناشناسی، پارسر

۱- این مقاله مستخرج از طرح پژوهشی با عنوان «شناسایی و تشریح تجربیات زبان‌شناسان معاصر روسیه در پردازش‌های ساختوازی، نحوی و معنایی برای استفاده در نظام‌های هوشمند» به شماره ۴۶۰۵۰۰۹/۱/۶ که با حمایت مالی معاونت پژوهشی دانشگاه تهران انجام شده است.

\* تلفن: ۶۱۱۱۹۰۶۹-۰۲۱، دورنگار: ۰۲۱-۸۸۶۳۴۵۰۰، E-mail: alreva@ut.ac.ir

## مقدمه

تفوق بر سد موانع زبانی، یکی از مهم‌ترین اهداف و وظایف توسعه فناوری ارتباطات است. ایده ساخت یک زبان مصنوعی ارتباطی، مانند اسپرانتو و یا سعی بر یک زبان کردن تبادل اطلاعات، هر کدام به عللی نتوانسته‌اند به موفقیت نایل شوند و به ناچار ایده اتوماسیون رونده ترجمه و بهره‌مندی از نظام‌های ترجمه ماشینی همچنان مورد توجه و مراجعه قرار دارد. (مارچوک ۱۷، ۱۹۹۰) در ایران نیز اخیراً دو مؤسسه تصمیم به تحقق این ایده داشته‌اند، نمونه‌هایی از نتایج کار آن‌ها نیز در این مقاله مورد بررسی قرار می‌گیرد.

در زبان‌شناسی معاصر به واسطه پیشرفت روزافزون در عرصه ترجمه ماشینی و احتیاج گسترده به این فناوری، ضرورت پرداخت به پژوهش‌های نوین در گردآوری 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، «چند معنایی، تعیین سبک و مشخص کردن حدود آن، از این‌گونه مشکلاتند، هم‌چنین آشکار است که برگرداندن جمله‌ای به زبان دیگر، لزوماً با تساوی نشانه‌های صرفی و زنجیری همراه نخواهد بود، نیز مفهوم کلمات نیز از زبانی به زبان دیگر، کاملاً همگن نخواهد بود». (بارخودووا (Barxodova)، ۱۲۲) برای مثال، فرض کنید فهرست بی‌پایان 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 را به دست دهد. او قادر است معانی مختلف یک جمله را تشخیص دهد، بگوید آیا جمله‌ای بی‌معناست یا بامعنا، و جمله‌های هم‌معنا را نیز در آن فهرست 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 را در هر یک از جمله‌های این فهرست نامتناهی 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 آن که معنی مورد نظر از طریق اطلاعات نحوی مربوط به آن تکواژ تعیین گردد. 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 و یا هر زبانی، در چنین کاربردهایی، باید شرح مقوله‌های واژگانی را نیز در برگیرد. بنابراین واضح است عدم تکمیل چنین توصیفاتی، اشکالاتی را در ترجمه‌های ماشینی تا مدت‌ها قبل از تکامل آن‌ها در بر خواهد داشت.

یادآور می‌شویم، 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 غیر ممکن تلقی کرده و تلاش در این زمینه را بی‌فایده دانسته‌اند. در واقع هنگامی که آنها بررسی نحو را آغاز کردند، روال‌های کشف که به نظر می‌رسید درباره آواها و واژه‌ها موثر افتاده‌اند، 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. (فالک، 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، همراه با افراط و تفریط بوده است. تجربه نشان می‌دهد نه پاک کردن صورت مسئله به سبک بلومفیلدیان و هم ساده‌انگاری‌هایی از استفاده و تعمیم‌های کلی روش‌هایی مانند چامسکی و دیگر ارائه‌کنندگان دستورهای جدید، کمکی به ارتقاء بهبود ترجمه ماشینی نتوانسته نکرده است. تاریخ کوتاه زبان‌شناسی کاربردی نشان می‌دهد، ساده‌انگاری از دستاوردهای تحولات انقلابی نه تنها کمکی به پیشرفت این علم نکرده، بلکه ارتقاء واقعی آن را به تأخیر انداخته است. این پروسه دقت و حوصله فراوانی را می‌طلبد و ساده‌لوحانه است، اگر تکمیل و تکامل آن را فقط منوط به ابداع نظریات جدید زبان‌شناسی بدانیم. به عنوان 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، به عنوان یک امکان منطقی و محتمل در زمینه ترجمه ماشینی تلقی کرده بود (دبیرمقدم، ۱۲۰). پیشرفت‌های بعدی در زبان‌شناسی نشان داد که این مسئله نسبت به آنچه که به نظر می‌رسید،

می‌بایست به طور 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 تنها یک فهم ساده از ساختار جملات نیست. (راسل، ۱۶)

چامسکی در کتاب خود چند اثر را نام برده است که دستور زایشی را نتیجه تلاش‌هایی برای استفاده از رایانه برای اهداف مختلفی همچون ترجمه ماشینی دانسته‌اند، ولی او این موضوع را رد می‌کند و ریشه‌های این دستور را در زبان‌شناسی سنتی می‌داند. (دبیر مقدم، ۱۷۹). با این حال چامسکی از ابراز دقیق انگیزه خود برای ارائه این دستور خودداری نموده است. 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 که به رغم تفاوت ظاهری و روئایی این زبان‌ها، در سطح جمله‌های هسته‌ای تشابه فراوانی بین آنها وجود دارد، وی ابراز امیدواری کرده بود که با رمزگشایی‌های همگانی جملات، زمینه برای ساخت الگوریتم‌های مناسب طبق روش‌های مرسوم شده در دستور گشتاری فراهم آید.

امروز به دلایل مختلف، که به برخی از آنها در این مقاله نیز اشاره می‌شود، معلوم شده است که این تصور عجولانه و تا حد زیادی خوشبینانه بوده است و عملاً چنین چیزی تحقق نیافته است. زیرا اشکالات در ترجمه ماشینی معمولاً آنچنان که 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 ژرف ساخت به رو ساخت است، از نظر کمی و کیفی تمام ساخت‌های نحوی را پوشش نمی‌دهند و هنوز هم استثنائات و نکات مبهم فراوانی 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 سازوکارهایی 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 غیر متخصص نیز می‌باشد، اگر چه چنین ارزیابی‌هایی دقیق و استاندارد نیست، اما می‌توان از آنها به عنوان قرینه و متمم تحقیقات تخصصی استفاده کرد.

### بحث و بررسی

امروزه گستره سازوکارهای ترجمه ماشینی به سرعت متنوع و دگرگون گشته‌اند و از حیطه متن مکتوب به گفتار زنده غیر هم‌زبانان توسعه یافته است. و با توسعه ابزارهایی مانند PDAها ارائه نرم‌افزار بر روی این دستگاه‌ها توسعه یافته و ارائه نرم‌افزارهای زبان طبیعی مانند تشخیص صوت یا حتی لغت‌نامه‌های مختلف بر روی این ابزارها افزایش زیادی یافته است. (پاسپلوا، ۵۸)

در ایران اولین کار مهم در گستره دانشگاهی و آکادمیک در زمینه ترجمه ماشینی پروژه شیراز بود (Shiraz Machine Translation) که توسط دانشگاه نیومکزیکو پشتیبانی می‌شد که آن هم در سال‌های ۱۹۹۱ به دلیل مسائل مالی و عدم تخصیص وام‌های مناسب متوقف ماند. این پروژه فارسی - انگلیسی از محدود پروژه‌هایی است که مستندات مناسبی را در این زمینه ارائه داده است، با این حال مورد استفاده همگانی قرار نگرفته است.

به بررسی نمونه‌ای از ترجمه انجام شده توسط نظام ماشینی پروژه شیراز می‌پردازیم. نمونه متنی که برای ترجمه از فارسی به انگلیسی مورد استفاده قرار گرفته است:

ابداعات علمی و مهندسی، موفقیت چشمگیر ما را افزایش می‌دهد. فناوری اطلاعاتی به تنهایی با ایجاد فرصت‌های مشاغلی که حقوق‌هایی نزدیک به ۸۰ درصد بیشتر از میانگین بخش خصوصی پرداخت می‌کنند، در حال حاضر یک سوم پیشرفت اقتصادی ما را به خود اختصاص می‌دهد. بار دیگر می‌بایست این را به خاطر سپرد که تحقیقات تامین شده توسط دولت، ابر کامپیوترها، اینترنت و ماهواره‌های ارتباطی را به وجود آوردند. به زودی محققان ابزارهایی را در اختیار ما قرار خواهند داد که قادر خواهند بود زبان‌های خارجی را به همان سرعت که می‌توانید صحبت کنید، ترجمه کنند، موادی ده برابر قوی‌تر از فولاد با وزنی بسیار کمتر از آن، و کامپیوترهای مولکولی به اندازه یک قطره اشک و با قدرت سریع‌ترین ابرکامپیوترهای کنونی.

ترجمه انجام شده توسط نظام ماشینی پروژه شیراز:

Science and engineering innovations are also propelling our remarkable prosperity. Information technology alone now accounts for a third of our economic growth, with jobs that pay almost 80 percent above the private sector average. Again, we should keep in mind: government-funded research brought supercomputers, the Internet, and communications satellites into being. Soon

researchers will bring us devices that can translate foreign languages as fast as you can speak; materials 10 times stronger than steel at a fraction of the weight; and molecular computers the size of a teardrop with the power of today's fastest supercomputers.

ارزیابی ترجمه از فارسی به انگلیسی توسط نظام ماشینی پروژه شیراز: به نظر نگارنده مقاله، ترجمه انجام شده قابل قبول و مفهوم است و ساختار نحوی جملات به طور نسبتاً مناسبی دارای شکل دستوری درستی است. این موضوع بیانگر دارا بودن سازوکارهای صحیح صرفی و نحوی و الگوریتم‌های دقیق در این سیستم است. به جمله اول دقت کنید.

Science and engineering innovations are also propelling our remarkable prosperity.

ابداعات علمی و مهندسی، موفقیت چشمگیر ما را افزایش می‌بخشند. ساختار صرفی کلمات درست و ساختار نحوی جمله نیز قابل قبول است. در ادامه، بررسی بخشی از متن ترجمه شده فارسی به انگلیسی را توسط نظام ترجمه ماشینی استایلوس (*stylus*) به روسی ترجمه کردیم و ارزیابی نگارنده این مقاله آن است، که ترجمه به دست آمده نیز مفهوم و قابل استفاده است.

Наука и технические новшества также продвигает наше замечательное процветание. Информационная технология, один теперь составляет третью часть(треть) нашего экономического роста, с рабочими местами, которые оплачивают почти 80 процентов выше среднего числа частного сектора. Снова, мы должны иметь в виду: финансируемое правительством исследование принесло суперкомпьютеры, Интернет, и спутники связи в являющийся. Скоро исследователи принесут нам устройства, которые могут перевести иностранные языки, с такой скоростью как Вы можете говорить; материалы 10 раз более сильно чем сталь во фракции(доле) веса; и молекулярные компьютеры размер teardrop со властью(мощью) сегодняшних самых быстрых суперкомпьютеров.

بنابراین می‌توان نتیجه گرفت از طریق این دو سازوکار، حتی امکان ترجمه برای روس‌زبان‌ها از فارسی به روسی نیز به طور غیر مستقیم مهیا شده است.

در این بخش از ارزیابی خود دو شرکت نسبتاً فعال در زمینه ترجمه ماشینی را که برخلاف پروژه دانشگاه شیراز محصول خود را به بازار نرم افزار ارائه کرده و مورد استفاده همگانی قرار گرفته‌اند را مورد بررسی قرار می‌دهیم.

### ۱- مترجم ماشینی دوران

سیستم مترجم ماشینی دوران محصول بخش هوش مصنوعی شرکت داده‌پردازان دوران است که به ادعای شرکت دوران با استفاده از آخرین دستاوردهای پردازش زبان طبیعی پیاده‌سازی شده است. هدف از این سیستم، ترجمه جملات انگلیسی به فارسی است. قابلیت‌هایی که شرکت دوران برای نرم‌افزار ابداعی خود عنوان نموده به شرح زیر است: این نرم‌افزار با استفاده از آخرین نوآوری‌ها و ایده‌های علمی مهندسی نرم‌افزار و با استفاده از آخرین دستاورد الگوریتم‌های پردازش زبان طبیعی تولید شده است. این شرکت ادعا کرده است که مترجم ماشینی دوران براساس شیوه انتقالی طراحی شده است که در آن از یک سیستم انتقالی که زبان مبدا را به زبان مقصد تبدیل می‌کند، استفاده می‌شود. در چنین سیستمی، زبان مبدا ابتدا وارد سیستم تحلیل شده و به یک ساختار میانی (که معمولاً ساختار گرامری جمله است) تبدیل می‌گردد، سپس با تعریف سیستم انتقال، این ساختار از زبان مبدا به ساختار معادل خود در زبان نهایی تبدیل می‌گردد و سرانجام از سیستمی دیگر برای ساخت جمله به زبان نهایی استفاده می‌شود. از ویژگی‌های این نرم‌افزار به موارد زیر دربروشور تبلیغاتی آن اشاره شده است:

- استفاده از معماری ترکیبی (Hybrid Model) حاصل از ایده‌های مبتنی بر قاعده و - - ایده‌های مدل‌های احتمالاتی.
- استفاده از مجموعه لغات وسیع شامل بیش از ۱۳۰/۰۰۰ کلمه انتقال یافته به زبان فارسی.
- امکان ترجمه جملات غیر نحوی و عناوین متون با استفاده از روش‌های تجزیه جزئی (Partial Parsing).

برای ارزیابی ادعای انجام شده متنی را به نظام ترجمه ماشینی دوران دادیم و ترجمه

۱- مراحل پارسینگ نظام ترجمه ماشینی دوران:



انجام شده نمونه‌ای برای ارزیابی ما است.

Dear Sirs,

As a result of requests from many of our clients we are eager to open a new information centre in St. Petersburg. Your recommendations proved to be really useful. Just imagine - to promote such an unordinary thing to your country means to be the first! Many had this chance, but didn't used it. Operating systems like that are now only in USA. But we can't keep moving further without your help, which means an invitation to visit our head office in Boston. We would be glad to see you between the first and seventh of August. Our cooperation looks very promising.

Sincerely

«سیرس» عزیز،

ما یک نتیجه درخواست‌ها از خیلی مشتری‌هایمان «اس» برای باز کردن یک «سنتر» اطلاعات جدید در «است» مشتاق هستیم. «پترسبارگ». توصیه‌هایتان واقعاً اثبات کردند مفید باشید. فقط تصور کردن - به کشورتان آنقدر یک چیز «آوردیناری» معنی می‌دهد باشد اولین ترفیع دادن! خیلی داشت این اتفاق افتادن اما آن را استفاده نکرد. سیستم‌های عملیاتی مانند آن در «بات» «یواس‌ای». به تنهایی حالا ما بیشترتر بدون کمک‌تان نتوانستن از «موینگ» نگهداری بکنیم کدام که یک دعوت با اداره «هید» مان در شهر بوستن ملاقات کردن را معنی می‌دهد. ما شاد هستیم شما را بین اولین و هفتم ماه اوت ببینیم. همکاری‌مان خیلی امیدوارکننده نگاه می‌کند.

«سینسرلی»

به نظر نگارنده این مقاله، علی‌رغم ادعاهای انجام شده، کیفیت ترجمه ضعیف و نتیجه قابل بهره‌برداری نیست. این نظام فعلاً احتیاج به ارتقاء در زمینه‌های نحوی، معنایی و حتی صرفی دارد. با دقت جزئی در متن ترجمه می‌توان پی برد، که این نظام در حال حاضر مانند لغت‌نامه کار می‌کند و از نظر بکارگیری الگوریتم‌های دستوری ضعیف و ناکارآمد است.

۲- مترجم مینا

براساس تبلیغات و اطلاعات ارائه شده توسط شرکت مینا سیستم، این مترجم ابزار مناسبی برای مراکز علمی، تحقیقاتی، آموزشی، مدیران، مترجمان، دانشجویان و کسانی که درک



سریع و ساده از مطلب انگلیسی را خواهانند، معرفی و تبلیغ شده است! برای بررسی مستند، چند عبارت و جمله نسبتاً ساده را که توسط مترجم مبنا انجام شده، مورد ارزیابی قرار می‌دهیم.

Financing of the health sector and health professionals whose expertise may lie in other areas

با سرمایه لازم را تأمین کردن از بخش بهداشتی و حرفه‌ای‌های بهداشتی که مهارتشان دروغ در منطقه‌های دیگر ممکن است.

A list of documents, which are distributed free of charge, appears before the contents page.

یک فهرست از سندها، که توزیع شده است مجانی، قبلاً صفحه محتویات ظاهر بشود.

staff development

توسعه کارکنان

systems and software development

سیستم و توسعه نرم‌افزار

tutions development

توسعه تأسیس

آنچنان‌که از ترجمه‌های صورت گرفته به دست می‌آید، کیفیت ترجمه سیستم مبنا حتی ضعیف‌تر از سیستم دوران است و این نظام دارای الگوریتم‌های ترجمه نحوی، معنایی و حتی صرفی نیست. این نظام صرفاً مانند لغت‌نامه کار می‌کند و در مثال‌های ارائه شده نیز فقط در حد کلمه و عبارت اضافی درست عمل کرده، ولی حتی در سطح ترجمه جملات ساده نیز ناکارآمد است، بنابراین ارائه آن به این شکل برای بازار نرم‌افزار عجولانه و باعث بی‌اعتمادی نسبت به این پدیده می‌شود.

غیر از سازوکارهای ماشینی فارسی به انگلیسی شیراز، انگلیسی به فارسی دوران و مبنا، سازوکار ماشینی فارسی به انگلیسی نسبتاً قوی و قابل قبولی به نام اختصاری اس ام تی اس (Language Weaver Persian SMTS Translation) در ایالات متحده آمریکا ساخته و به بهره‌برداری رسیده، که در دسترس همگانی قرار گرفته است. این نظام توسط یک شرکت آمریکایی با همکاری دانشگاه کالیفرنیا جنوبی جدیداً زبان فارسی را به مجموعه زبان‌های مترجم ماشینی خودش اضافه کرده است، و دارای اسم تجاری Language Weaver است.

ارزیابی ترجمه از فارسی به انگلیسی نظام ماشینی اس ام تی اس

متن فارسی برای ترجمه:

انقلاب اسلامی ایران بیش از هر چیزی مرهون فرهنگ اسلامی و تعلیم حیات بخش اسلام و تداوم و گسترش چنین انقلابی در گرو بقا و رشد فرهنگ اسلامی در نسل‌های موجود و آینده است. این وظیفه بیش از همه بر عهده نهادهای آموزشی و پرورشی، به خصوص دانشگاه‌هاست.

الف) ترجمه ماشینی توسط نظام ماشینی اس ام تی اس:

Iran's Islamic revolution than any other due to Islamic culture and education section of Islam and the continuation and expansion of such a revolutionary in the future survival and development in Islamic culture and the future generations. It was the duty of all the responsibility of educational institutions and upbringing, especially college quadrangle.

ارزیابی نگارنده این مقاله آن است، که ترجمه به دست آمده مفهوم و قابل استفاده است و ارزش ویراستاری را دارد.

ب) ترجمه انسانی برای مقایسه با ترجمه ماشینی نظام ماشینی اس ام تی اس:

Iran's Islamic revolution more than anything else is grateful for the Islamic culture and teaching of Islam; the continuation and expansion of such Islamic revolution is guaranteed by survival and development of Islamic culture in future generations. This burden, more than anything else, is the duty of educational institutions especially universities.

در ادامه ارزیابی و بررسی خود از نظام ترجمه ماشینی Language Weaver، جملات ساده و مرکبی را با ساختارهای نحوی آسان، متوسط و مشکل را توسط این نظام ترجمه کردیم تا به طور مقایسه‌ای و عینی کیفیت دستوری جملات را آشکار کنیم. به مواردی از ترجمه آن که با نرم‌افزار گیت (Gate) دانشگاه شفیلد از نظر ساختار نحوی مورد ارزیابی قرار گرفته است، توجه کنیم. منظور ما از این بررسی ارزیابی ماشینی ساخت جملات در محل گره‌های نحوی آنهاست.

My name is Alireza Valipour.

نام من علیرضا ولی پور است.

ТИРЕ, :4  
прил. сущ. иностр. лекс.  
Мое имя - Alireza Valipour .



همچنین می‌توان با ملاحظه دقیق‌تر پی برد که داده پردازی در سطح نحو نیز علی‌رغم اعمال الگاریتم، حالت اضافی در عبارت‌هایی مانند:

Scientific foreign languages University of Tehran

و یا

Graduated from Moscow State University

چندان موفق نبوده است. اگر چه جمله قابل درک بود، و قابل ویرایش است. این سیستم<sup>۱</sup> علی‌رغم سرعت بالای خود چون فقط از طریق اینترنت قابل بهره‌برداری است، مقداری افت سرعت پیدا می‌کند. (در حدود صفحه‌ای ۱۰ ثانیه).

در ادامه ارزیابی خود و برای مقایسه سازوکارهای روسی با غربی، یکی از این سازوکارها را که در دسترس عمومی است، به نام استایلووس مورد ارزیابی و محک قرار می‌دهیم. برای این منظور صحت ترجمه‌های نظام ترجمه ماشینی استایلووس (stylus) از روسی به زبان‌های آلمانی، فرانسه و انگلیسی را توسط اساتید دانشکده زبان‌های خارجی دانشگاه تهران مورد ارزیابی قرار دادیم.

متن اصلی به زبان روسی که از آن به زبان‌های آلمانی، فرانسه و انگلیسی ترجمه ماشینی صورت گرفته است.

## ЗАЯВЛЕНИЯ

Заявления подразделяются на следующие категории: затребованные и незатребованные. Затребованное прошение пишется в ответ на объявление ("Предлагаемые места", "Вакантные места", "Предлагаемая работа") или по просьбе предполагаемого нанимателя. Незатребованное заявление адресуется фирме с надеждою, что там есть или будет вакансия. Как правило, незатребованные заявления посылаются по совету третьего лица.

От лиц, подающих заявление, требуется представить на рассмотрение автобиографию или листок личных данных. Автобиография содержит данные

۱- نمونه آزمایشی سیستم پس از مکاتبه برای زمان ۳۰ روزه قابل دریافت است، اگر چه مسئولان این شرکت به مکاتباتی که با آی‌پی رایانه‌های ایرانی باشد، متأسفانه پاسخ مثبت نمی‌دهند!

о претенденте, о воспитании и квалификации. В листке личных данных информация о претенденте сведена в таблицу. Предполагаемые наниматели обычно хотят посмотреть дипломы претендентов, рекомендательные письма. Копии этих документов могут сопутствовать заявлению, так как посылка оригиналов была бы слишком рискованна. Многие фирмы посылают формы заявления кандидатам с просьбой заполнить и вернуть их с автобиографией и копиями необходимых личных документов.

سؤالات ارائه شده به استادان<sup>۱</sup> و پاسخ‌های ایشان با توجه به ترجمه انجام شده، به شرح زیر بوده است.

ترجمه و ارزیابی ترجمه توسط استاد گروه آلمانی  
ترجمه از روسی به آلمانی:

### **Die Anträge**

Die Anträge werden auf die folgenden Kategorien unterteilt: angefordert und uneingelost. Der angeforderte Antrag wird als Antwort auf die Erklärung (Anzeige) ("die angebotenen Stellen", "die unbesetzten Stellen", "die angebotene Arbeit") oder auf Bitte von dem vermuteten Mieter geschrieben. Der uneingeloste Antrag wird an die Firma mit der Hoffnung adressiert, daß dort existiert oder es wird die Vakanz. In der Regel, werden die uneingelosten Anträge auf Rat von dem Dritter gesandt.

Von den Personen (Gesichter), die den Antrag reichen, ist es erforderlich, die Autobiographie oder das Blatt (Schein) der persönlichen Daten vorzulegen. Die Autobiographie enthält die Daten über den Bewerber, über die Erziehung und die Qualifikation. Im Blatt der persönlichen Daten ist die Information über den Bewerber in die Tabelle gebracht. Die vermuteten Mieter wollen die Diplome der Bewerber, der Empfehlungsschreiben gewöhnlich durchsehen. Die Kopien dieser Dokumente können (dürfen) den Antrag begleiten, da die Sendung der Originale

۱- استادان محترمی که از گروه‌های آلمانی، فرانسه و انگلیسی به تحقق این ارزیابی کمک نمودند به ترتیب عبارتند از: سرکار خانم دکتر بهجت - آقای دکتر اسفندی و سرکار خانم دکتر معرفت.

allzu riskiert ware. Viele Firmen senden die Formen der Anträge den Kandidaten mit Bitte, sie mit der Autobiographie und die Kopien der notwendigen persönlichen Dokumente auszufüllen und zurückzugeben.

- ۱- (س) این متن که توسط ترجمه ماشینی از زبان روسی به آلمانی ترجمه شده، تا چه حد مفهوم است؟ درصد تقریبی موفقیت ترجمه را اعلام نمایید. لطف کرده ترجمه به فارسی این متن را نیز مرفوم فرمایید.
- (ج) - چیزی حدود ۹۵٪ از مفهوم منتقل شده است.
- ۲- (س) به نظر جنابعالی این ترجمه در حد دانشجویان کدام سال می‌تواند باشد؟
- (ج) - در حد یک دانشجوی خوب سال چهارم
- ۳- (س) سبک انتخاب لغات و گرامر به کار رفته تا چه حدی قابل قبول است؟
- (ج) - گرامر خوب و قابل قبول است و با توجه به این که در مواردی چند کلمه معادل با هم پیشنهاد شده‌اند باید متن را ویرایش کرد.
- ۴- (س) اگر نظر شخصی دیگری دارید عنوان فرمایید.
- (ضمناً قابل ذکر است که مدت صرف شده توسط برنامه برای ترجمه فقط پنج (۵) ثانیه بوده است.)

(ج) - پاراگراف اول کمی نامفهوم است و من ترجیح دادم موارد تردید آفرین را در یک فرهنگ لغت ابتدا کنترل کنم و سپس ترجمه نهایی را انجام دهم، مع هذا با توجه به این که متن از روسی توسط ماشین مترجم به آلمانی ترجمه شده است و یک ماشین فاقد بسیاری از شرایط لازم برای انجام یک ترجمه خوب است و مدت زمان انجام این کار هم طبق آنچه عنوان شده ۵ ثانیه بیشتر نبوده است، عمل ترجمه بسیار خوب انجام شده است.

ارزیابی ترجمه توسط استاد محترم گروه انگلیسی  
ترجمه از روسی به انگلیسی:

#### THE APPLICATIONS

The applications are subdivided into the following categories: demanded and undemanded. The demanded application is written in reply to the announcement ("Offered places", "Vacant places", "Offered work") or at the request of the prospective employer. The undemanded application is addressed to firm with hope,

that there is or there will be a vacancy. As a rule, the undemanded applications are sent on advice of the third person. From the persons giving the application, it is required to submit the curriculum vitae or leaf of the personal data. The curriculum vitae contains the data on the applicant, about education and qualification. In a leaf of the personal data the information on the applicant is shown in the table. The prospective employers usually want to see (overlook) the diplomas of the applicants, letters of recommendation. The copies of these documents can accompany the application, as the parcel (sending) of the originals would be too risky. Many firms send the forms of the applications to the candidates with the request to fill and to return them with the curriculum vitae and copies of the necessary personal documents.

- ۱- (س) این متن که به وسیله ترجمه ماشینی از زبان روسی به انگلیسی ترجمه شده تا چه حد مفهوم است؟ درصد تقریبی موفقیت ترجمه را اعلام نمایید.
- (ج) - البته بستگی به خواننده هم دارد ولی اگر خواننده مشکل زبانی نداشته باشد، کاملاً قابل فهم است.
- ۲- (س) به نظر جنابعالی این ترجمه در حد دانشجویان کدام سال می‌تواند باشد؟
- (ج) - دانشجوی ترم چهارم به بالا.
- ۳- (س) سبک انتخاب لغات و گرامر به کار رفته تا چه حدی قابل قبول است؟
- (ج) - کاملاً قابل قبول و قابل فهم است، اما خواننده نیز باید به طور فعال در فهم آن کمک کند.
- ۴- (س) اگر نظر شخصی دیگری دارید عنوان فرمایید.
- (ج) - فکر می‌کنم یک لغت کلیدی در متن نیست که باعث می‌شود اول خواننده متن را تا آخر بخواند و بعد متوجه شود مربوط به چیست.

ارزیابی ترجمه توسط استاد گروه فرانسه  
ترجمه از روسی به فرانسه:

### LES DEMANDES

Les demandes (declaration) sont subdivisees en categories suivantes: demande

et non demande. La demande est écrite en réponse à l'annonce ("les Emplois à pourvoir", "les places Vacantes", "L'OFFRE d'emploi") ou sur la demande de l'employeur éventuel. La demande (déclaration) non demandée est adressée à la société dans l'espoir que la ktre ou il y aura une vacance. En General, les demandes (déclarations) non demandées sont envoyées sur les conseils du tiers. Des personnes posant la candidature, il faut présenter le curriculum vitae ou le feuillet des données personnelles. Le curriculum vitae contient les données sur le candidat, sur l'éducation et la qualification. Dans le feuillet des données personnelles l'information sur le candidat est réduite à la table. Les employeurs éventuels veulent examiner ordinairement les diplômes des candidats, les lettres de recommandation. Les copies de ces documents peuvent accompagner la demande (déclaration), puisque l'envoi des originaux était trop risqué. Plusieurs sociétés envoient les formulaires aux candidats avec la demande les de remplir et rendre avec le curriculum vitae et les copies des pièces d'identité nécessaires.

- ۱- (س) این متن تا چه حد مفهوم است؟ حدوداً درصد آن مشخص فرمایید  
(ج) - متن کاملاً مفهوم است و بنظر من تا ۹۸٪ ترجمه موفق بوده است.
  - ۲- (س) به نظر جنابعالی این ترجمه که توسط برنامه ترجمه ماشینی صورت گرفته است در حد دانشجویان کدام ترم رشته زبان فرانسه می‌تواند باشد؟ لطف نموده ترجمه فارسی آن را مرقوم فرمایید.  
(ج) - در حد دانشجوی سال چهارم این ترجمه ارائه شده است.
  - ۳- (س) سبک انتخاب لغات و گرامر به کار رفته تا چه حدی قابل قبول است؟  
(ج) - کاملاً قابل قبول و قابل فهم است و متن ترجمه مطلوب است.
  - ۴- (س) اگر نظر شخصی دیگری دارید عنوان فرمایید.  
(ج) - انتخاب در اغلب موارد درست بود، فقط چند مورد جابجایی واژه وجود دارد و می‌توان گفت ارائه متن از نظر گرامری در پاراگراف دوم بهتر است.
- در ادامه این پژوهش مقایسه‌ای، همان متنی را که با سازوکار ماشینی دوران ترجمه کردیم و پس از ترجمه مورد ارزیابی قرار دادیم، به سازوکار روسی ماشینی استایلوس داده و سپس بدون آن‌که موضوع ماشینی بودن ترجمه را فاش کنیم از چند تن از استادان و ویراستاران زبان



روسی خواستیم، که مورد ارزیابی قرار دهند و در صورت لازم تصحیح نمایند.  
متن انگلیسی برای ترجمه:

*Dear Sirs,*

As a result of requests from many of our clients we are eager to open a new information centre in St.Petersburg. Your recommendations proved to be really useful. Just imagine - to promote such an unordinary thing to your country means to be the first! Many had this chance, but didn't used it. Operating systems like that are now only in USA. But we can't keep moving further without your help, which means an invitation to visit our head office in Boston. We would be glad to see you between the first and seventh of August. Our cooperation looks very promising.

*Sincerely*

ترجمه از انگلیسی به روسی توسط سازوکار ماشینی استایلوس:

*Уважаемые господа,*

В результате запросов от многих из наших клиентов мы стремимся открывать новые информационные центр в С.-Петербурге. Ваши рекомендации оказалось, были действительно полезными. Только вообразите - продвигать такую необычную вещь на вашу страну, означает быть первым! Многие имели этот случае, но не использовали это. Операционные системы подобно этому - теперь только в США. Но мы не можем продолжать перемещаться далее без ваше помощи, что означает приглашение посетить наше главное бюро в Бостоне. Мы были бы довольны видеть Вас между первым и седьмым августа. Наше сотрудничество выглядит очень многообещающим.

*Искренне*

ارزیابی ترجمه توسط نگارنده مقاله:

ترجمه انجام شده کاملاً مفهوم و سبک و مقولات صرفی و نحوی بسیار خوب از زبان مبداء به زبان مقصد انتقال داده شده است. با این حال ۱۴ مورد اشکال غیر بنیادی در ترجمه

به چشم می‌خورد که زیر آن خط کشیده‌ایم. این اشکالات در محدودهٔ نحو، سبک و انتخاب غیر دقیق کلمات است و در واقع شکل دستوری کلمات از نظر صرفی آن‌ها کاملاً درست بوده و خللی در معنای کلی متن ایجاد نمی‌کند. با این حال نگارندهٔ این مقاله ترجیح نداد، که ترجمهٔ برتر انسانی را خود ارائه نماید و ترجمهٔ گویشگران متخصص زبان روسی را برای مذاقه و ارزیابی دقیق‌تر اخذ و ارائه نموده است.

ترجمه‌های از انگلیسی به روسی توسط اساتید روسی:

ترجمهٔ اول توسط خانم پروفیسور ناتالیا اسلوخای استاد دانشکدهٔ زبان‌های خارجی دانشگاه شوچنکوو کیف:

В результате запросов от многих наших клиентов, мы стремимся открыть новый информационный центр в С.-Петербурге. Ваши рекомендации, оказалось, были действительно полезными. Только вообразите - продвигать такую необычную вещь на вашу страну, означает быть первым! Многие имели этот случай, но не использовали его. Операционные системы подобно этой - теперь только в США. Но мы не можем продолжать перемещаться далее без вашей помощи, что означает приглашение посетить наше главное бюро в Бостоне. Мы были бы рады видеть Вас с первого по седьмое августа. Наше сотрудничество будет очень многообещающим.

Искренне

ترجمهٔ دوم توسط خانم پروفیسور ناتالیا یورونا استاد دانشکده زبان‌های خارجی دانشگاه

مینسک:

Доктор Валипур, мне кажется, первое предложение лучше написать так:<sup>1</sup>  
«В связи с большим количеством запросов, поступающих от наших клиентов, мы собираемся открыть новый информационный центр в Санкт-Петербурге».

Второе предложение... Может быть, там должно быть «Ваши рекомендации оказались бы по-настоящему полезными»?

۱- آنچنان‌که ملاحظه می‌شود، این استاد (Наталья Юревна) تغییرات در ترجمه خود را بیشتر روی سبک کلمات متمرکز کرده است.

И еще, мне кажется, вместо «Подобные операционные системы – теперь только в США» правильное было бы сказать «Подобные операционные системы в настоящее время распространены только в США». «Главное бюро» - правильное «главный офис». «Надеемся на дальнейшее сотрудничество» - это традиционная фраза в конце делового письма... «С уважением» - традиционная подпись в конце делового письма... Я, например, подписываюсь так:

С уважением,

Наталья Юревна.

ترجمه سوم توسط پروفسور یوری نیکلایوویچ مارچوک، استاد دانشگاه دولتی مسکو و ویراستار مجله علمی - پژوهشی دانشکده ادبیات.

Уважаемые господа!

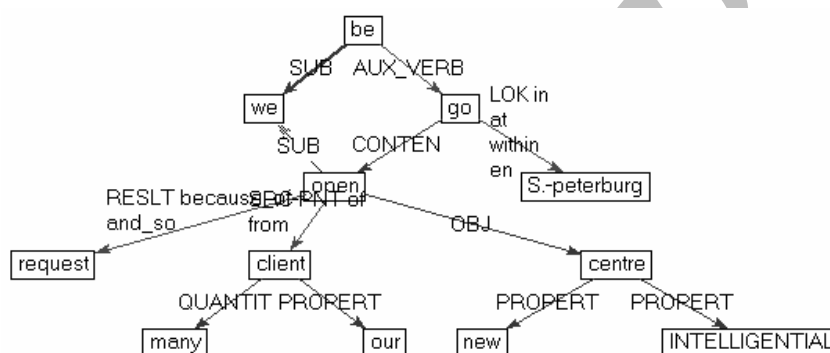
В результате запросов от многих наших клиентов мы собираемся открыть новый информационный центр в С.-Петербурге. Ваши рекомендации, как оказалось, были действительно полезными. Только представьте - продвигать такую необычную вещь в вашей стране означает быть первым! Многие имели этот шанс, но не использовали его. Подобные операционные системы – теперь только в США. Но мы не можем продолжать двигаться дальше без вашей помощи, поэтому приглашаем Вас посетить наше главное бюро в Бостоне. Мы были бы рады видеть Вас между первым и седьмым августа. Надеемся на дальнейшее сотрудничество.

С уважением

از مقایسه این ترجمه‌ها می‌توان به این نتیجه رسید، که علی‌رغم تفاوت در نحو و انتخاب واژگان، معنا و مفهوم جملات با یکدیگر مشابه و تعارضی با ترجمه ماشینی ارائه شده نخست نیز ندارد.

از آنجایی که تحلیل معنایی جملات کار ساده‌ای نیست و خود احتیاج به پژوهشی جداگانه، دارد به این مقوله در این مقاله پرداخته نمی‌شود، و در این قسمت از پژوهش خود

جمله ترجمه شده بوسیله سازوکار ماشینی استایلوس را توسط تحلیل گر نحوی دیالینگ<sup>۱</sup> مورد ارزیابی قرار می‌دهیم.<sup>۲</sup> این تحلیلگر ماشینی ابتدا الگوریتمی برای ساخت جمله زبان مبدا براساس قواعد سازه‌ای، ژرف‌ساخت‌های آن و سرانجام روساخت به وجود می‌آورد و سپس با استفاده از ژرف‌ساخت همگانی آن، ژرف‌ساخت و روساخت جمله مقصد را ارایه می‌کند. به نمودار جمله زبان مبدا توجه کنید.

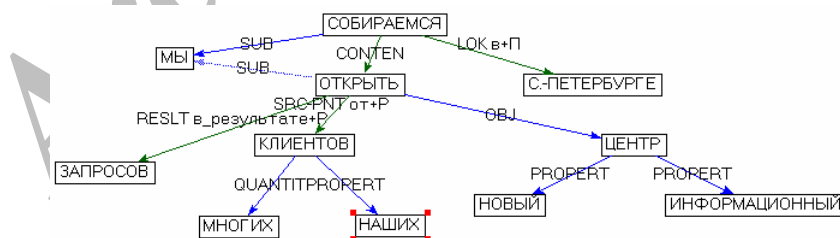


اولین جمله عبارت است از:

В результате запросов от многих наших клиентов мы собираемся открыть новый информационный центр в С.-Петербурге.

برای آن‌که تجزیه جمله ساده‌تر شود، جمله‌واره‌ای را که در آغاز جمله آمده است، حذف کرده‌ایم. در نتیجه، جمله اصلی عبارت خواهد بود از:

мы собираемся открыть новый информационный центр в С.-Петербурге.



#### 1- Диалинг

۲- دیالینگ به اعتقاد کارشناسان و زبان‌شناسان شاغل در حیطه‌های کاربردی و به خصوص ترجمه ماشینی، بهترین نرم‌افزار موجود در روسیه است، که در سازوکارهای ماشینی بعنوان مغز تحلیلگر آن‌ها مورد استفاده قرار می‌گیرد.

همان‌طور که ملاحظه می‌شود، این کنترل تایید می‌کند، ساختار نحوی جملات در ترجمه با سازوکار ماشینی استایلوس منطبق با الگوریتم‌های تحلیل گر نحوی بوده و قابل قبول است. لازم به یادآوری است، این کنترل و مقایسه در رابطه با جملات دیگر نیز انجام شد و همین نتیجه قابل قبول را در بر داشت.

### نتیجه

به واسطه پیشرفت روزافزون در عرصه ترجمه ماشینی و احتیاج گسترده به این فناوری، ضرورت پرداخت به پژوهش‌های نوین در جمع‌آوری و پردازش داده‌های زبانی، اجتناب‌ناپذیر می‌نماید. سابقه ساخت سیستم‌های ترجمه ماشینی امروزه به بیش از پنجاه سال می‌رسد. ولی علی‌رغم پیشرفت‌های فراوان در سطح تطابق‌های واژگانی و صرفی، هنوز پیشرفت چندانی در حیطه‌های نحوی و معناشناسی 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 هم‌اند، و زبان‌ها فقط در گزینش تأویل با هم اختلاف دارند. این موضوع هنوز به علت ابهاماتی که در عمل بروز می‌کند، به صورت 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 برای ترجمه به دست آید.

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 که پاک کردن صورت مسئله به سبک بلومفیلدیان و هم ساده‌انگاری‌هایی از استفاده و تعمیم‌های کلی روش‌هایی مانند چامسکی و دیگر ارائه‌کنندگان دستورهای جدید کمکی به ارتقاء بهبود ترجمه ماشینی نتوانسته نماید.

بنابراین برای ارتقای سازوکارهای 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 گستره است.

تجربه ما در این پژوهش نیز تأیید این موضوع بود، که اگر سازوکار ماشینی هر چه بیشتر امکان فراهم شده از الگوریتم‌های نحوی برایش فراهم شده باشد، ترجمه‌های موفق‌تری ارائه خواهد داد. مصداق عملی این ادعا در عملکرد و ارزیابی ما از سازوکار ترجمه ماشینی استایلوس مشاهده شد.

### کتاب‌شناسی

فالک، جولیا اس. (۱۳۷۷). *زبان‌شناسی و زبان*. مترجم خسرو غلامعلی‌زاده. چاپ پنجم. موسسه چاپ و انتشارات آستان قدس رضوی

دبیر مقدم، محمد. (۱۳۸۳). *زبان‌شناسی نظری*. تهران: نشر سمت.

راسل، رابرت. (۱۳۸۳). *هوش مصنوعی*. مترجم رامین رهنمون. تهران: نشر ناقوس

Барходова Т.И. (1995). *Новая лексикография*. М.: ВЦП,—183 с.

Марчук Ю. Н. (1990). *О моделировании перевода*. Вопросы информационной теории и практики. М.: ВИНТИ. № 43.

Поспелова Д.А. (2003). *Искусственный интеллект*. М.: ВЦП, 2003.—259 с.

Archive of SID