**НЕЙРОНДЫҚ МАШИНАЛЫҚ АУДАРМАҒА АРНАЛҒАН ПОСТРЕДАКЦИЯЛЫҚ ТӘСІЛДЕРІН ЗЕРТТЕУ**

***Асанғали Айгерім Серікбайқызы***

*2-курс магистранты,*

*Ақпараттық технологиялар кафедрасы,*

*әл-Фараби атындағы Қазақ ұлттық университеті,*

*Қазақстан, Алматы қ.*

***Калиева Айдана Канатқызы***

*2-курс магистранты, «Программалық жобаларды басқару (МИФИ)» мамандығы,*

*Ақпараттық технологиялар кафедрасы,*

*әл-Фараби атындағы Қазақ ұлттық университеті,*

*Қазақстан, Алматы қ.*

***Саурбек Улжан Болатханқызы***

*2-курс магистранты,*

*Ақпараттық технологиялар кафедрасы,*

*әл-Фараби атындағы Қазақ ұлттық университеті,*

*Қазақстан, Алматы қ.*

***Тукеев Уалишер Ануарбекович***

*проф., ғылыми жетекші*

*Ақпараттық Жүйелер мамандығы,*

*Ақпараттық Технологиялар кафедрасы,*

*әл-Фараби атындағы Қазақ ұлттық университеті,*

*Қазақстан, Алматы қ.*

**АННОТАЦИЯ**

Бүгінгі таңда аудармашыларға көмектесетін бірқатар жүйелер бар. Нейрондық машиналық аударманы постредакциялау қарапайым және қайталанатын мәтіндерді, нұсқаулықтарды аудару қажет болған кезде ең жақсы таңдау болып табылады. Машиналық аударма компанияларға қаржылық және уақыттық шектеулерге байланысты бұрын аударылмаған ақпараттық материалдарды аударуға мүмкіндік береді. МА бойынша технологиялық шешімдерді жұмыс көлеміне және тапсырыс берушінің кез-келген талаптарына, соның ішінде арнайы жазу стилі мен грамматикалық ережелерге, компания қолданатын нақты терминология мен сөздіктерге, сандар мен күндер форматына арналған арнайы ережелерге және материалдардағы басқа да көптеген мәліметтерге бейімдеуге болады.

**Кілттік сөздер:** машиналық аударма, постредакциялау, аударма көрсеткіштері, нейрондық желі.

**НМА постредакциялау: теориялық аспектілері**

МА жүйелерінің айтарлықтай дамуына қарамастан, оның аудармасы көбінесе маманның түзетуін қажет етеді. Бұл процесс машиналық аударманы постредакциялау (РЕМТ) деп аталады және алдын-ала белгіленген талаптарға сәйкес машиналық аударма мәтініне түзетулер енгізуді білдіреді. Мәтінді постредакциялау редакциялаудан ерекшеленеді, бірінші жағдайда бастапқы мәтін – бұл машина жасаған аударма, ал екіншісінде – адам. Постредакциялау кезеңінен өткен мәтін кейіннен редакторға стилистикалық, грамматикалық және лексикалық қателерді түзету үшін берілуі мүмкін. Редакциялау маманнан, аудармашыдан немесе постредактордан гөрі жоғары шеберлікті талап етеді [1].

Статистика бойынша МА-мен өзара әрекеттесу кәсіби аудармашы жұмысының маңызды құрамдас бөліктерінің бірі болып отыр. Еуропаның бірқатар кәсіби бірлестіктері (Elia GALA, EUTAC, EMT, LIND) әзірлеген 2018 жылы аударма саласының дамуы жайлы есебінде 2018 жыл бірінші кезекте есте қалатыны атап өтілді, себебі, дәл осы жылы Еуропаның аударма компанияларының да, фрилансер аудармашыларының да 50%-дан астамы МА-ны пайдаланатынын растады. Биылғы жылы МА-ны постредакциялау одан да танымал болды: еуропалық аударма компанияларының 37% - ы өз тәжірибесінде РЕМТ-ны пайдалану үлесінің артқаны туралы хабарлайды, ал 17% - ы жақын арада осы қызмет түрімен айналысуға ниет білдірді [2].

Дегенмен, МА-ны постредакциялаумен бірге қолдану аудармашының жылдамдығы мен тиімділігінің артуына әкеледі деп айтуға болмайды. Нәтиже көбінесе МА-ның сапасына байланысты. Зерттеулер көрсеткендей, жоғары сапалы МА-ны постредакциялау кәсіби аудармашылардың жылдамдығы мен тиімділігін арттыра алады. Алайда, төмен сапалы МА-ны постредакциялау кері нәтижелерді көрсетеді. Сонымен қатар, МА-ны постредакциялау тиімділігінің тағы бір маңызды факторы – постредакторды дайындау болып табылады, ол сонымен қатар постредактордың негізгі тәжірибелерін стандарттауды және әр түрлі деңгейдегі түпкілікті өнімнің сапасына қойылатын талаптарды қамтиды. Сондықтан, МА-ны постредакциялаудың тиімділігін арттыру үшін келесі сұрақтарға жауап беру керек: жоғары сапалы мәтінді шығару үшін мәтінді машиналық аудармаға қалай дайындауға болады? Тиімді жұмыс істеу үшін постредакторға қандай ұсыныстар беру керек?

Жоғарыда айтылғандай, постредакциялау тиімділігі МА-ның аудару нәтижесінің сапасына тікелей байланысты. Төмен сапалы аударманы постредакциялау аудармашының басынан бастап жасаған аудармасына қарағанда ұзақ уақыт алуы мүмкін. МА-ның аудару нәтижесінің сапасына "аударма көрсеткіштері" немесе "теріс аударма көрсеткіштері" деп аталатын бастапқы мәтіннің белгілі бір сипаттамалары әсер етеді. Бұл сипаттамаларды зерттеудің көп бөлігі ағылшын тілінде жүргізілгеніне қарамастан, алынған нәтижелер барлық тілдерге қолданыла алады.

Бұл сипаттамалардың біріншісі – сөйлемнің ұзындығы мен құрылымы. Ұзын сөйлемдер де және өте қысқа сөйлемдер де аудару нәтижесінің сапасының төмендеуіне әкелуі мүмкін. МА жүйесі арқылы көбінесе ұзын сөйлемге дұрыс талдау жасау қиынға соғады, ол мағынаны дәл жеткізу үшін өте маңызды. Ал қысқа сөйлемдер шектеулі контекстке байланысты мағынасыз болуы мүмкін. Аударма үшін толық емес сөйлемдер, яғни бір немесе екі негізгі сөйлем мүшесі болмайтын, сондай-ақ салалас мәнді жалғаулықтары бар сөйлемдерді аудару қиын. Сөйлемнің тағы бір аудармалылығының теріс көрсеткіші – бұл сөз тіркестерінің болуы, мысалы, тиісті атаулар, күрделі зат есімдермен бекітілген комбинациялар және т.б.

Постредакциялаудан кейінгі уақытты қысқарту үшін алдымен бастапқы мәтінмен жұмыс істеу керек (алдын-ала редакциялау). Бұл процесс бастапқы мәтіннің грамматикалық, орфографиялық және логикалық дұрыстығын тексеруді қамтиды; сөз тіркесі мен сөйлем деңгейінде сөз ретін өзгерту; мағынаны бір мәнді жеткізу үшін идиомалық өрнектерді ауыстыру; күрделі тыныс белгілерін өзгерту. Сонымен қатар, MА жүйесін дұрыс орнату аударманың сапасын жақсарта алады. Ереже негізіндегі машиналық аударма — Rule-based Machine Translation (RBMT) жағдайында мұндай күйге келтіруге тақырыптық сөздіктерді пайдалану арқылы, ал статистикалық және гибридтік жүйелер жағдайында терминологиялық сәйкестікті қамтамасыз ету үшін бастапқы деректерді сүзу арқылы қол жеткізіледі.

МА-ны постредакциялау тиімділігі көбінесе постредакторға байланысты. Аударма және постредакциялау саласындағы әртүрлі тәжірибесі бар субъектілердің қатысуымен жүргізілген зерттеулер барысында олардың арасындағы айтарлықтай айырмашылықтары анықталды. Біріншіден, пост-редакциялау жылдамдығындағы айырмашылық байқалды. Бұған бірнеше факторлар себеп болуы мүмкін: мәтінмен жұмыс жасаудың жалпы дағдылары, постредакциялау тәжірибесі және постредакциялау құралдарымен жұмыс істеу. Сондай-ақ, аударма тәжірибесі мен постредакциялау жылдамдығы арасында оң байланыс жоқ екенін атап өткен жөн. Сонымен, лингвист Ганс Петер Крингс тәжірибелі аудармашылар бастаушы аудармашыларға қарағанда постредакциялауға көп уақыт жұмсайтынын көрсетті. Неғұрлым өзекті зерттеуде тәжірибелі аудармашылардың жаңадан келгендерге қарағанда артықшылықтарын анықтау мүмкін болмады. Алайда, сынақтар барысында постредакциялаудың тиімділігінің аса маңызды критерийлері –постредакциялауға арналған ұсынымдарды сақтау және аударма мақсатына сәйкес сапаға бағдарлану анықталды [2,3,4].

Бұл кезеңде постредакциялау анықтамасына оралу маңызды: алдын-ала белгіленген талаптарға сәйкес машиналық аударма мәтініне түзетулер енгізу. Бұл кейбір жағдайларда постредактордың мақсаты түпкілікті өнімнің идеалды сапасына емес, қолайлы сапаға қол жеткізу дегенді білдіреді. Мысалы, пайдаланушылар жасаған мазмұнның аудармасы грамматикалық және стилистикалық тұрғыдан дұрыс болмауы керек, ол тек бастапқы хабарламаның мағынасын дәл жеткізуі керек. Басқа жағдайларда жоғары сапалы аударма қажет болуы мүмкін, оның деңгейі кәсіби маманның аудармасына сәйкес келуі керек.

Постредакциялау процесін тиімдірек және белгілі бір сапаға бағыттау үшін аударма сапасының әр түрлі деңгейіне жалпы талаптарды және осы талаптарды орындау үшін стандартталған тәжірибелерді әзірлеу қажет. Бұл тапсырманы TAUS аналитикалық орталығы (Translation Automation User Society) жүзеге асырды, онда постредакторларға да, тапсырыс берушілерге де түпкілікті өнімге нақты талаптар қоюға және соның негізінде техникалық тапсырма жасауға көмектесетін постредакциялау нұсқауларының жиынтығы дайындалды. Құжатта машиналық аударманы постредакциялау сапасының екі деңгейі нақты анықталған.

Сапаның бірінші деңгейі – "қолайлы" сапа ("good enough") бастапқы хабарламаның мағынасын дәл жеткізетін аудармаға сәйкес келеді, бірақ сонымен бірге грамматикалық, синтаксистік және стилистикалық тұрғыдан дұрыс болуы міндетті емес. Мәтін компьютер аударғандай көрінуі мүмкін. "Қолайлы" сапаға қол жеткізу үшін постредактор аударманың семантика тұрғысынан дұрыс екендігіне көз жеткізуі керек, оның кездейсоқ қоспалары немесе кемшіліктері жоқ, сонымен қатар мәдениеті жағынан мазмұн жоқ.

Сапаның екінші деңгейі, "жарияланымға қойылатын талаптарға жауап беретін" ("publishable quality") сапасы алушыға бастапқы хабарламаның мағынасын дәл жеткізуге және толық түсінікті болуға, сондай-ақ стиль, грамматика және синтаксис нормаларына сәйкес келуге тиіс аудармаға сәйкес келеді. Яғни, кәсіби маман аудармасына жақын болу [5].

Машиналық аударманы постредакциялау сапасына қойылатын осы талаптар ISO 18587: 2017 аударма қызметтерін ұсыну жөніндегі халықаралық стандартта бекітілген. Құжатта taus аналитикалық орталығы жасаған екі деңгейге ұқсас сапаның екі деңгейі (жеңіл және толық постредакциялау) көрсетілген.

Бүгінгі таңда машиналық аударма технологияларын жетілдіру, сондай-ақ аударма жылдамдығын арттыру қажеттілігі постредакциялау аударма қызметінің көптеген аспектілерінде кәсіби маман аудармасына балама бола бастайды. Алайда, постредакциялаудың қарапайымдылығына қарамастан, бұл процестің тиімділігі көбінесе МА-ның бастапқы мәтінінің сапасына, сондай-ақ алдын-ала бекітілген постредакциялау тәжірибесіне және түпкілікті өнімнің сапасына қойылатын талаптарға байланысты. Көптеген адамдардың санасында әлі күнге дейін қауіп төндіретін машиналық аударма кәсіби аудармашы жұмысының маңызды компоненттерінің біріне айналады. Дегенмен, компьютерлік жүйелер адамды толығымен алмастыра алмайды. Сондықтан машиналық аударманы постредакциялау тақырыбы лингвистикалық зерттеулер саласында өзекті болып қала береді, бұл одан әрі аударма процесінің тиімділігі мен аударма сапасына әсер етуі тиіс[6].

**Әдебиеттер тізімі:**

1. *Mikel L.F*. Making sense of neural machine translation // Translation Spaces 6:2 (2017) 291–309, DOI 10.1075/ts.6.2.06for.
2. *He W., Wu Hua., Wang H.* Improved neural machine translation with SMT features // Thirtieth AAAI conference on artificial intelligence, 2016.
3. *Luong M., Manning C.* Achieving open vocabulary neural machine translation with hybrid word-character models, 2016. arXiv preprint.
4. Bertoldi, N., Cettolo, M., Federico, M., 2013. Cache-based online adaptation for machine translation enhanced computer assisted translation. In:Machine Translation Summit, Nice, France, September.
5. Blain, F., Schwenk, H., Senellart, J., 2012. Incremental adaptation using translation information and post-editing analysis. In: International Workshopon Spoken Language Translation, Hong-Kong, China, pp. 234–241.
6. Antonio L. Lagardaa, Daniel Ortiz-Martínez, Vicent Alabaub, Francisco Casacuberta, 2014, Translating without in-domain corpus: Machine translationpost-editing with online learning techniques