

ХАЛИМАГ ХЭЛНИЙ ҮНДЭСНИЙ БИЧВЭРИЙН САН БҮРДҮҮЛСЭН ҮГЗҮЙН ТАЙЛБАРЛАЛЫН ТУРШЛАГА

Ванькаева А. Э. (Россия)

1. Оршил

Энэхүү нийтлэлд халимаг хэлний хэлзүйн толь бичиг бүтээх ажлын үндсэн зарчмуудыг өгүүллээ. Хэлзүйн толь бичиг нь хэлний үгзүйн бүтэн мэдээллийг агуулдаг. Үүнийг хамгийн алдартай нь «Грамматический словарь (словоизменение)» А.А.Зализняка 1977 юм. Үг өөрчлөх, хувиргах иймэрхүү үгзүй мэдээлэл байгаа нь халимаг хэлний үндэсний бичвэрийн сан гэх мэт төрөл бүрийн хайгуулын тогтолцоо байгуулах нөхцөл боломжийг бий болгож байна.

ОХУ-ын ШУА-ийн гүйцэтгэх зөвлөлийн үндсэн хөтөлбөрийн хүрээнд 2011 онд байгуулсан Халимаг Хэлний Үндэсний Бичвэрийн Сан (ХХҮБС) нь дотроо 20-21р зууны эхэн үеийн утга зохиолын ойролцоогоор 800 мянга үгтэй бичвэр эхүүдийг багтаасан байдаг. Хэлзүйн толь бичгийн үгийн сангийн хэмжээ нь ойролцоогоор 23 мянган үгтэй. Үгзүйн автомат анализийн тогтолцоог Т. Архангельский боловсруулж, Corpus Technologies компани дорнод армяны үндэсний бичвэрийн сангийн хайлтын суурь програмыг өгчээ.

2. Хэлзүйг толь бичиг

Хэлзүйн тольбичигт Муниевийн 1977 гаргасан халимаг-орос хэлний үгийн баазыг ашигласан.

2.1. Толь бичиг бүтээх ажлын зорилго

Халимаг хэлний хэлзүйн толь бичиг зохиох ажлын явцад гарсан хамгийн чухал даалгавар нь үгийн баазад орсон бүх лексемүүдийн хувирах хэлбэрүүдийг тодорхойлох байсан. Хэлзүйн (үгзүйн) хувьд ижил хэлбэртэй олон хэлбэрүүдийг хувирсан үгийн анги (словоизменительным класс) гэдэг. Ингээд лексем бүрт тохирсон хувирсан үгийн анги зооно.

Үүнд нэгдүгээрт халимаг хэлний үг хувирах хэм хэмжээ, заагийг нь тодорхойлох нь чухал. Энэ нь тэр болгон тодорхой байдаггүй. Хоёрдугаарт, халимаг хэлэнд тохирсон бүх хэлзүйн айнуудыг харгалзан үзэх шаардлагатай өөрөөр хэлбэл хэлзүйн айнуудыг системчлэх хэрэгтэй. Эцэст нь бүх хувилбаруудыг хувирах хэв шинжээр нь (үгийн үндсийн эгшигээс шалтгаалж өөр өөр залгавар авах гэх мэт) бүлэглэж хүснэгтэнд оруулах хэрэгтэй.

Мөн түүнчлэн халимаг хэлний утга зохиолын хэлний зөв бичих дүрэмтэй холбоотой онцлогуудаар нь ангилах хэрэгтэй.

Дээр дурьдсан зүйлүүдийг харгалзан үзэж Сегалович, Маслов нарын 1999 онд гаргасантай төстэй хэлзүйн толь бичиг зохиох дүрэм журам алгоритм боловсруулсан.

Энэхүү журамын үндсэн хэсгүүд:

1. Парадигмууд (хувилбарууд)-ын хүснэгтийн зүүн хэсэгт байрлах хэлзүйн айн өгөгдлүүдийн жагсаалтыг гаргах. Энэ жагсаалтанд орсон өгөгдлүүд нь Лейпцигийн дүрмийн дагуу эрэмблэгдэж, халимаг хэл шинжлэлийн сонгодог бүтжээлүүдэд [Санжеев1983],[Харчевникова1996] заасан дүрэмд нийцүүлсэн. Мөн түүнчлэн [Биткеев и др. 1983] болон [Сай 2009] –аас материал ашигласан болно.

2. Үндэс, үгсийн сангийн хамгийн бага утгат хэсэг ба лексемийн үндэсийн дугаар гэх мэт заалтуудын жагсаалт. Лексемууд нэгээс дээш тооны үндэстэй байж болно (хан - хаан 'хан'), сэлгэлттэй (күн 'человек' – эмтн 'люди') байж болно. Лексемийн үг хувиргах ангийн заалтуудаар тухайн үгийн төрөл, хэлбэрийг тодорхойлж болно. Сингармонический үзүүлэлтүүд нь тухайн үгэнд ямар сингармонический нөхцөл дагавар шаардлагатайг заана.

3. Грам (нэр болон үйл үг үүсгэх дагаварууд) –уудын жагсаалт

Лексем бүрд тохирсон бүх бүтээврүүдийг тодорхойлсоны дараа байж болох бүх үндсийн хувилбаруудыг олно. Дараа нь энэ хувилбар бүрийг хүснэгтийн нэг нэг нүдэнд оруулна. Ингэхдээ нэг нүдэнд нэгээс дээш үндэс орсон байж болно (например, арслцг+д | арслц+д). Энэ нь зөв бичих зүйн дүрэмтэй холбоотой.

Эцэст нь хүснэгтийн нүдэнд байгаа үгийн үндсэн дээр жагсаалтаас дагаваруудыг холбоно.

2.2. Үг хувиргах/бүтээх анги үүсгэх

Тодорхой нэг лексемд тохирсон үг хувиргах/бүтэх анги үүсгэх дээрхи үйл явц нь хоёр үеэс бүрднэ:

- Дүрс тэмдэгтүүдийн жагсаалтын задлан дүгнэлт

Үгийн бааз (словник) –д орох бүх замыг нягталж шалгасан: үг бүтээх бүтээлч дагаваруудын тусламжтайгаар үүссэн үгийн анги, бүлгийг тодорхойлно (например, -лнн, -лт (нэр үг үүсгэх дагавар), -ч (үйлдэгчийн утга заасан дагавар) гэх мэт). Энэ үе шатанд олон тооны хэлбэрийг оруулаагүй.

- үгийн сангийн утгаар нь үг бүтээх/хувиргах хэлбэрийг тодорхойлох үе. Энэ дэд үед үгийн утга талаас нь авч үзнэ. Жишээ нь халимаг хэлний нэрийн сонгодог нэгдүгээр ялгал нь (-н төгссөн) өөрөө дотроо хоёр хуваагддаг юм байна.

1. олон тооны хэлбэрт –д залгадаг (модн 'дерево' - модд 'деревья'.

яман 'коза' -ямад 'козы');

2. олон тоон дээр араасаа –муд авдаг (эшелон 'эшелон' -эшелонмуд 'эшелоны').
Энэ бүлэгт голдуу зээлдмэл үгс орно.

Ийм аргаар бүтээсэн хэлзүйн толь бичиг нь 30 гаруй нэрийн, 4 үйлийн дэд бүлгүүдтэй болж байна. Нэрийн хэсэгт 150 орчим нүд, харин үйлийн хэсэгт үйлт нэр, нөхцөл үйл үгийг тооцвол ойролцоогоор мянга болно.

3. Үгзүйн задлан дүгнэлтийн асуудал

Үгзүйн задлан шинжилгээ хийх явцад халимаг хэлний задлага шинжтэй холбоотой хүндрэл гарч байсан.

3.1. Үг бүтээх гинжин хэлхээ

Хоёрдугаар үйлдүүлэх хэв (в терминологи [Куликов 1994]) (шат-а-лн-хтореть' CAUS-CAUS-INF)-ийг авч үзэх үед үйлдүүлэх хэвийн дагавар нь сэлгэх эсвэл хос болох асуудлыг шийдэж чадаагүй. Мөн 2-р тийн ялгал (нана-н-аг 'тетя' GEN-ACC)-ын асуудлыг ч шийдэж чадаагүй. Ийм төрлийн асуудлууд нь үг хувирах ёс бус харин үг бүтэх ёстой холбоотой нь илэрхий бөгөөд ийм үг бүтэх гинжин хэлхээг хайлтын механизмд тусгах шаардлагатай.

3.2. Үгийн сан-хэлзүйн ижил нэр

Үгийн сан-хэлзүйн ижил нэр гэдэг нь нэг ижил (например, слово ширэ 'стол' в именительном ивинительном падеже) эсвэл өөр өөр (ке - прилагательное 'нарядный' и глагол 'делать' в форме императива) лексемийн хэлзүйн утга давхацахыг хэлнэ.

Одоогоор үгийн сан-хэлзүйн ижил нэрийг гараар эсвэл автоматаар авах гэсэн хоёр арга байна. Автоматаар авахад хэд хэдэн хүндрэл гарч байна. 1-рт, ихэнхи программуудад зөв оноолт өгнө гэсэн баталгаа байхгүй. 2-т, халимаг хэлэнд тохирсон програм байхгүй. Иймээс үгийн сан-хэлзүйн ижил нэрийг ХХҮБС-аас гараар хийх аргыг сонгож байна.

“САНЖЕЕВСКИЕ ЧТЕНИЯ-7”(2012 он) сэтгүүлээс орчуулав.

*НХУС-ийн Монгол хэлний тэнхимийн багш,
Л.Жумдаан доктор(Ph.D), дэд профессор,*