УДК 519.765/681.3

## КОМПЬЮТЕРДИК ЛИНГВИСТИКАНЫН МАСЕЛЕЛЕРИ

***Жумабаева А.Н.***, *Н. Исанов атындагы КМКТАУ КЛМАБ кафедрасынын ага окутуучусу, Кыргызстан, aikymbat@mail.ru.*

*Компьютердик лингвистика тексттик билдирүүнү автоматтык түрдө иштеп чыгуу системасында негизги ролду ойнойт. Компьютердик лингвистиканын негизги илимий маселеси тексттин негизинин маанисин түшүндүрүүнү моделдөө жана сөздү синтездөө болуп эсептелет.*

***Түйүн сөздөр***: *Компьютердик лингвистика, тексттик маалымат, машиналык сөздүктөр, издөө системдери, машиналык котормо, жасалма интеллект, табигый тил.*

## ПРОБЛЕМ КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ

***Жумабаева А.Н.***, *старший преподаватель каф. КЛиМК КГУСТА им.Н.Исанова, Кыргызстан, aikymbat@mail.ru*

*В системах автоматической обработки текстовой информации важную роль играет компьютерная лингвистика. Центральными научными проблемами компьютерной лингвистики являются проблема моделирования процесса понимания смысла текстов и проблема синтеза речи.*

***Ключевые слова***: *компьютерная лингвистика, текстовая информация, машинные словари, машинный перевод, искусственный интеллект, естественные языки.*

## PROBLEMS OF COMPUTER LINGUISTICS

Zhumabaeva A.N.**,** *senior teacher of dept. CLICC KSUCTA n.a.N.Isanov, Kyrgyzstan, aikymbat@mail.ru*

*The computer linguistics plays an important role the system of the automatic processing to text information. The central scientific problems of the computer linguistics are a problem of modeling of the process of the understanding the meaning text and problem of the syntheses speech.*

***Key words***: *computer linguistics, textual information, mashine dictionary, mashine translation, artificial intellect, natural languages.*

XX-кылымдын ортосунда ЭЭМдин түзүлүшү жана кибернетикалык идеялардын тез өнүгүшү мурун элестетүүгө мүмкүн болбогон жаңы илимдердин пайда болушуна

себеп болду. Алар көп учурда бири бири менен байланышта болбогон илимдердин кесилишинде пайда болду. Биология жана инженердик илимдердин кесилишинде бионика, психология жана лингвистиканын кесилишинде психолингвистика, ошондой эле эсептөө техника менен лингвистиканын кесилишинде жаңы илим пайда болду. Бул жаңы илим бир нече жолу аталышын өзгөрткөн. Ал математикалык лингвистика, андан кийин, түзүлүштүк лингвистика, эсептөө лингвистикасы болуп өзгөргөн. Акырында компьютердик лингвистика деп азыркы аты бекем аталып калды. Жаңы илимдин пайда болушуна эки себеп шартталган. Биринчиден, лингвист-изилдөөчүлөр лингвистикага жетишпеген тактыкты заманбап так илимдердин (баарынан мурда математика) жардам менен тактыкка ээ болууга жардам берет деп үмүттөнүшкөн. ЭЭМдин пайда болушу ушул үмүттөрдү күчөттү, анткени көпчүлүк тилчилерге башынан эле компьютер бул бир гана «тез иштөөчү арифмометр» эмес, ошондой эле тексттер менен иштөөдө автоматташтыруучу күчтүү курал экендиги далиденди. Көп эмгек, убакыт талап кылынуучу процесстерди автоматташтыруу мүмкүнчүлүгү пайда болду. Мисалы, тексттерди статистикалык иштетүү, ар түрдүү сөздүктөрдү жана лексикалык картотекаларды тейлөө ж.б.у.с. иштер. Экинчиден, компьютерлердин пайда болушу менен дээрлик колдонуучулардын даярдыгы жок болгондуктан компьютерлерди иштетүү маселеси пайда болду. 70-жылдардын башында компьютердик лингвистика боюнча атайын жыйнактар жана журналдар чыгарыла баштады, лингвистика жана жасалма интеллект боюнча конференцияларда тийиштүү лекциялар окула баштады, акырында бул илимдин маселерине аралган дүйнөлүк форумдар өткөрүлдү.

Компьютердик лингвистика бул табигый тил менен берилген маалыматтарды автоматтык түрдө иштеп чыгаруу менен байланышкан билим тармагы. Компьютердик лингвистиканын борбордук илимий маселелерине: тексттердин маанисин түшүнүү процессин моделдөө маселеси (тексттен анын маанисинин формалдуу түрүнө өткөрүү), кеп синтезинин маселеси (формалдык түрдө берилген маанини табигый тилдеги тексттерге өткөрүү) кирет. Бир катар колдонмо маселелерди чечүүдө, атап айтканда, компьютерге тексттерди киргизүүдө автоматтык түрдө аныктоо жана каталарды оңдоо маселеси, автоматтык түрдө бир тилден экинчи тилге тексттерди которууда, компьютер менен табигый тилде байланыштырууда, текст документтерин автоматтык түрдө классификациялоо жана индекстөөдө, аларды автоматташтырып жалпылоо, толук текстүү маалымат базаларынан документтерди издөөдө бул маселелер пайда болду.

Компьютердик лингвистикада түзүлгөн жана колдонулган линвистикалык каражаттарды шарттуу түрдө эки бөлүккө бөлсө болот: декларативдик жана процедуралык. Декларативдик бөлүккө тил жана кеп бирдиги, сөздүктөр, тексттер жана ар кандай түрдөгү грамматикалык таблицалар, ал эми процедуралык бөлүккө тексттерди жана грамматикалык таблицаларды, тилдин жана кеп бирдиктерин башкаруу каражаттары кирет.

Компьютердик лингвистиканын колдонмо маселелерин чечүүдөгү ийгилик баарынан мурда декларативдик каражаттардын компьютердин эс тутумуна толук жана так берилишинен көз каранды. Азыркы күндө бардык өнүккөн мамлекеттерде компьютердик лингвистика тармагында иш чаралары жүрүп жатса да бул маселелер

керектүү деңгээлде чечиле элек. Ошентсе да, компьютердик лингвистика тармагындагы олуттуу илимий жана практикалык жетишкендиктерди белгилей кетсе болот. Кээ бир өлкөлөрдө (Россия, США, Япония ж.б.) бир тилден экинчи тилге тексттерди машинанын жардамы менен которуунун эксперименталдык жана өндүрүштүк системдери түзүлгөн.

Табигый тилде компьютер менен байланыш үчүн бир катар эксперименталдык системдер түзүлгөн, терминологиялык маалымат банктарды, тезуарустарды, эки тилдүү жана көп тилдүү машина сөздүктөрүн (Россия, США, Германия, Франция ж.б.) түзүү боюнча иш чаралары жүргүзүлүүдө, кепти автоматтык талдоо жана синтездөө системдер курулууда (Россия, США, Япония ж.б.), табигый тилдердин моделдердин түзүү тармагында изилдөөлөр жүргүзүлүп жатат.

Колдонмо компьютердик лингвистиканын негизги методикалык маселеси болуп, тексттик маалыматты автоматтык түрдө иштетүү системдердин декларативдик жана процедуралык компоненттеринин ортосундагы катнашка туура баа берүү кирет. Эмнеге артыкчылык берүү керек: грамматикалык жана семантикалык маалыматтарга бай, салыштырмалуу анча чоң эмес сөздүк системдерге таянган күчтүү эсептик процедураларгабы же салыштырмалуу жөнөкөй процедуралык каражаттар менен күчтүү декларативдик компонентергеби? Адистердин пикирлери боюнча экинчи жолу артыгыраак болуп эсептелет, анткени ал практикалык максатка батыраак жеткирет.

Маалыматтарды чогултуу, сактоо, кайра иштетүү жана издөө процесстеринин лингвистикалык камсыздоо компьютердик лингвистиканын алдында биринчи кезектеги маселеси болуп эсептелинет. Эң маанилүүлөрүнө төмөнкүлөр кирет:

1. Машиналык сөздүктөрдүн лингвистикалык иштетүү жана түзүүнү автоматташтыруу;
2. тексттерди компьютерге киргизүүдөгү каталарды табуу жана оңдоо процессин автоматташтыруу;
3. маалыматтык суроо-талаптарды жана документтерди автоматтык түрдө индекстөө;
4. автоматтык түрдө документтерди классификациялоо жана жалпылоо;
5. маалыматтарды бир тилдик жана көп тилдик маалымат базалардан издөө процесстерин лингвистикалык камсыздоо;
6. тексттерди бир тилден экинчи тилге машинанын жардамы менен которуу;
7. колдонууучулардын автоматташтырылган интеллектуалдуу маалымат системдери менен байланышты камсыз кылган лингвистикалык процесстерди куруу;
8. факторгафиялык маалыматтарды формалдаштырылбаган тексттерден бөлүп алуу.

Тексттик маалыматтарды атоматтык түрдө иштетүүчү системдердин дээрлик баарында машиналык сөздүктөр алардын ажырагыс бир бөлүгү болуп саналат. Алар туруктуу илимий техникалык түшүнүктөрдү билдирүүчү сөз же сөз айкалыш сөздүктөрү болушу мүмкүн. Сөздүктөрдү түзүүдө текстеттердин лексикалык түзүлүүнүн жогорку даражада чагылдырууга умтулуу керек. Автоматташтыруу

каражаттарын кеңири колдонуунун негизинде мындай жумуштарды акылга сыярлык мөөнөттө аткарууга мүмкүн.

Маалымат борборлордун практикалык иш аракеттеринде тексттерди компьютерге киргизүүдө каталарды табуу жана оңдоонун автоматташтыруу маселесин чечүү зарылдыгы бар. Бул комплекстүү маселе шарттуу түрдө үч маселеге: тексттерди орфографиялык, синтаксикалык жана семантикалык текшерүү маселелерине бөлүүгө болот. Биринчи маселе, жетишерлик күчтүү машиналык сөздүктөрдү колдонулуучу морфологиялык талдоо процедурасы менен чечилиши мүмкүн.

Синтаксистик текшерүү маселеси орфографиялык текшерүү маселесине караганда бир кыйла оор. Биринчиден, ал өзүнүн түзүлүшүнө милдеттүү компонент катары орфографиялык текшерүүлөрдү да камтыйт, экинчиден, формалдаштырылган тексттерде синтаксикалык талдоо маселеси толук көлөмдө чечиле элек. Ошондой болсо да, тексттерди жарым-жартылай синтаксистик талдоо жүргүзүүгө болот.

Мааниси жагынан кеткен каталарды аныктоо максатында тексттерди семантикалык текшерүү меселесин жасалма интеллект маселелер классына киргизүүгө болот. Ал бир гана адамдын ой жүгүртүү процессин моделдөөнүн негизинде толук көлөмдө чечилиши мүмкүн.

Маалымат издөө сапатынын маанилүү мүнөздөмөсү болуп, анын толуктугу жана тактыгы болуп саналат. Издөө толуктугу тил менен кеп (сөз жана сөз айкалышы) бирдигинин ортосундагы парадигмалык байланышты максималдуу эске алуу жолу менен, ал эми тактык алардын синтагматикалык байланышты эске алуу жолу менен камсыздалат. Издөөнүн толуктугу менен тактыгы тескери байланышта болушат деген пикир бар, ушул эки мүнөздөмөнүн бирин жакшыртуу чаралары экинчинин начарлашына алып келет. Бирок бул издөөнүн туруктуу белгиленген логикасына гана тиешелүү. Эгер бул логиканы өркүндөтсө, анда бул эки мүнөздөмө бирдей жакшырат.

Компьютердик лингвистиканын маанилүү жана перспективдүү маселеси колдонуучу интеллектуалдуу автоматташтырылган маалымат системдери менен табигый тилде же табигый тилге жакын тил менен байланышты камсыздоо лингвистикалык процесстердин түзүлүшү болуп эсептелинет. Заманбап интеллектуалдуу системдерде маалымат формалдаштырылган түрдө сакталгандыктан, лингвистикалык процесстер адам менен компьютер ортосунда ортомчулук ролун аткаруу менен бирге төмөнкү негизги маселелерди чечиш керек: 1) табигый тилде киргизилүүчү маалыматтык суроо-талаптар менен билдирүү тексттеринен алардын формалдык тилге которуу (маалыматты ЭЭМге киргизүүдө); 2) алынган билдирүүлөрдүн маанисин формалдаштырылган түрдөн анын табигый тилге өткөрүү. Биринчи маселе киргизилүүчү суроо-талаптар менен билдирүүлөрдү морфологиялык, синтаксистик жана концептуалдуу талдоо аркылуу, ал эми экинчи алынган билдирүүлөрдү концептуалдык, синтаксистик жана морфологиялык синтез жолу менен чечилиши керек.

Компьютердик лингвистиканын жарым кылымдык өнүгүү тажрыйбасы анын маселелерин конструктивдүү алгоритмдик жол менен чечүүдөгү кеңири мүмкүнчүлүктөрдү көргөздү. Ошону менен бирге алгоритмдик жол менен чечүүдө

чектөөлөр бар экенин аныктады. Кээ бир татаал жагдайларда алгоритмдик жол менен чечүү эффективдүү эмес экен, бул учурларда аналогия ыкмасын колдонуу жакшы.

Компьютердик лингвистика тармагында жарым кылымда бир кыйла илимий жана практикалык натыйжалар: тексттерди бир тилден башка тилдерге машиналык котормо системдери, тексттерде маалыматтар издөө системдери, ооз эки кепти автоматтык түрдө талдоо жана синтездөө системдери жана башкалар иштелип чыгарылган. Бирок буларды иштеп чыгарууда көп тоскоолдуктар жана кыйынчылыктар болду.

Кыргызстан да компьютердик лингвистикасынын маселелерин кароодо четте калган жок. Азыркы учурда кыргыз тилинде компьютердик лингвиситка маселелерин бир катар лингвист окумуштуулар кароодо. Т.С. Сыдыковдун 30 жылдан ашуун бул тармакта эмгектенип келген, анын салымы абдан чоң. Азыркы учурда биздин мамлекетте компьютердик лингвистика тармагындагы адистерди даярдоо процесси да жүрүп жатат. Компьютердик лингвистиканын маалымат технологиясынын өнүгүүсүндө чоң роль ойногонун айтып кетүү керек. Маалыматтык коом түзүү доорунда бул илим эң негизги факторлордон болуп калат.

**Колдонулган адабияттар:**

1. Чеповский А. Неразрешимые проблемы компьютерной лингвистики //Компьютера, 2002 №30, с 3-7.
2. Ермаков А.Е. Компьютерная лингвистика. 2002. 380 с.
3. Садыков Т.,Жумалиева Г.Э., Түмөнбаева М.Ж., Шаршембаев Б. “Кыргыз тилинин компьютердик лингвистикасынын негиздери”. Бишкек.: Имак Офсет, 2015. - 400 б. ISBN 978-9967-27-897-4
4. Белоногов Г.Г., Калинин Ю.П., Хорошилов А.А. Компьютерная лингвистика и перспективные информационные технологии. М., 2005. 268 с.

**Рецензент:** к.ф.н., доцент Жумалиева Г. Э.

УДК 82.0:82-95

## КОТОРМО ТЕОРИЯСЫНДАГЫ ТРАНСФОРМАЦИЯЛЫК КОНЦЕПЦИЯЛАР

***Жумалиева Г.Э.,*** *филол.и.к., доцент Н.Исанов атындагы КМКТАУнун ЖМТИси, Кыргызстан,* *gulira_1972@mail.ru*

***Куттубек к. Чынара*** *Ж. Баласагын атындагы КУУнун ККФсынын магистранты, Кыргызстан,kuttubekova.chynara@mail.ru*

*Бул макалада теория котормосунун негизги концепциялары каралган.*

***Түйүн сөздөр:*** *которуу, тил, концепциялар, сапат*