

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Державний заклад

«Південноукраїнський педагогічний університет ім. К.Д. Ушинського»

кафедра перекладу і теоретичної та прикладної лінгвістики

С.А. Любимова

«Проблеми автоматизованого перекладу»

(Конспект лекцій)

для здобувачів вищої освіти за другим (магістерським) рівнем за

спеціальністю: 035 Філологія (германські мови і літератури

(переклад включно)

Рекомендовано до друку рішенням кафедри перекладу і теоретичної та прикладної лінгвістики Державного закладу «Південноукраїнський національний педагогічний університет імені К. Д. Ушинського» (протокол № від 04 листопада 2024р.)

Любимова С. А., доктор філологічних наук, доцент, доцент кафедри перекладу і теоретичної та прикладної лінгвістики. Проблеми автоматизованого перекладу (конспект лекцій для здобувачів вищої освіти за другим (магістерським) рівнем за спеціальністю: 035 Філологія (германські мови і літератури (переклад включно): Одеса: Університет Ушинського, 2024. 49 с.

Рецензенти:

Юмрукуз А.А., кандидат філологічних наук, доцент, завідувачка кафедри германських і східних мов та методики їх навчання

Демчук А.І., кандидат філологічних наук, доцент, доцент кафедри германської філології та методики викладання іноземних мов

Системи машинного перекладу надають такі переваги, як оперативність, практичність, доступність, гнучкість, багатofункціональність. Програма-перекладач – це інструмент, який дає змогу підвищити ефективність роботи перекладача у разі його грамотного застосування. Конспект лекцій з навчальної дисципліни «Проблеми автоматизованого перекладу» спрямований на ознайомлення студентів із принципами і методами розробки систем машинного перекладу, їх лінгвістичним забезпеченням, специфікою такого перекладу і проблемами, пов'язаними з машинним перекладом. Конспект лекцій призначений для здобувачів вищої освіти за другим (магістерським) рівнем за спеціальністю: 035 Філологія (германські мови і літератури (переклад включно). Конспект складається із 8 лекцій, зміст і послідовність яких відповідає навчальній програмі дисципліни, схваленій комісією з якості освіти Державного закладу «Південноукраїнський національний педагогічний університет імені К. Д. Ушинського» 2023 р.

ЗМІСТ

Лекція 1.	Види автоматизованого перекладу	4
Лекція 2.	Алгоритм діяльності в процесі машинного перекладу	8
Лекція 3.	Особливості автоматизованого машинного перекладу	12
Лекція 4.	Основні принципи декодування мовного сигналу	19
Лекція 5.	Багатозначність як проблема автоматизованого перекладу	24
Лекція 6.	Помилки машинного перекладу писемних текстів та способи їх виправлення	31
Лекція 7.	Розпізнавання письмових текстів. Членування мовного потоку	38
Лекція 8.	Розпізнавання усних висловлювань. Особливості фонематичного декодування сигналів	45

Лекція 1.

Види автоматизованого перекладу

- 1) Поняття «машинний переклад».
- 2) Види машинного перекладу.
- 3) Автоматичні і автоматизовані системи машинного перекладу.

Машинний переклад (МП) – це процедура перекладу тексту з однієї мови на іншу без участі людини за допомогою комп'ютерної програми. На фундаментальному рівні МП виконує просту заміну слів тексту вихідною мовою на слова цільової мови.

Корпусні методи дають змогу здійснювати складніші переклади: вони краще працюють із фонетичною типологією, перекладом ідіом, а також визначенням нетипових конструкцій. Наразі системи МП не можуть працювати так само, як перекладач-людина, але в найближчому майбутньому це може стати реальністю.

Для опису перекладацького інструментарію в технології перекладу зазвичай використовують такі терміни:

- машинний переклад;
- людський переклад зі застосуванням програмного забезпечення;
- машинний переклад за участю людини;
- повністю автоматичний високоякісний (машинний) переклад.

Види МП. Можна виокремити чотири основні типи машинного перекладу:

1. Статистичний машинний переклад.
2. Машинний переклад на основі правил.
3. Гібридний машинний переклад.
4. Нейронний машинний переклад.

Статистичний МП – це такий вид перекладу, який працює на основі статистичних моделей, які будуються на основі вивчення величезних обсягів контенту двома мовами – оригіналу й перекладу. Статистичний МП визначає відповідності між словами вихідної мови та словами цільової мови.

Для базових завдань можна вибрати цей вид машинного перекладу. Типові помилки, однак, не дають змоги очікувати якісного результату. Найсуттєвіший недолік – те, що статистичний МП не враховує контекст і тому може регулярно помилятися.

МП на основі правил. Щоб створити переклад речення, ця модель МП проводить граматичний аналіз вихідного тексту й будує переклад з урахуванням відповідних граматичних правил. Результат такого перекладу потрібно суттєво редагувати.

Гібридний МП. Цей вид машинного перекладу є сумішшю статистичного МП і МП на основі правил. У цій моделі використовується пам'ять перекладів, завдяки чому результат якісніший. Проте навіть гібридний МП має недоліки й тому потребує великого обсягу редагування, тобто залучення перекладачів або редакторів.

Нейронний МП. У цьому типі машинного перекладу для побудови статистичних моделей із метою перекладу використовуються нейромережеві моделі (за аналогією з роботою людського мозку). Суттєвою перевагою нейронного МП є те, що він надає єдину систему для роботи з вихідним і цільовим текстами.

Автоматизований переклад (англ. Computer – Aided Translation) – переклад текстів на комп'ютері з використанням комп'ютерних технологій. Від машинного перекладу останній відрізняється тим, що увесь процес перекладу здійснюється людиною, комп'ютер лише допомагає їй зробити готовий текст або за менший час, або кращої якості. На сьогодні існує кілька

класифікацій систем машинного перекладу, до огляду яких ми й звертаємося.

Отже, сучасні системи машинного перекладу (СМП) підрозділяються на системи прямого типу, трансферні системи перекладу та системи машинного перекладу семантичного типу.

Системи прямого перекладу – найчисельніші, оскільки почали створюватися ще в 1950-60-і роки для фіксованих пар мов. У цих системах словник і синтаксис вхідної мови аналізуються лише в тій мірі, в якій це необхідно для ідентифікації правильних виразів вихідної мови та порядку розташування слів. Системи прямого типу нині практично не використовуються, ставши надбанням історії розвитку СМП.

Системи перекладу з використанням мови-посередника (семантичні), служать для відображення «сенсу» вхідного тексту, який перетворюється в семантичні і синтаксичні уявлення, загальні для декількох вихідних мов. Цей метод застосовується зазвичай при необхідності перекладу вихідного тексту на кілька мов.

Системи перекладу з трансфером більш складні, ніж попередні типи, оскільки мови-посередники застосовуються двічі – перший раз при перекладі з вхідної мови, другий – при перекладі на вихідну мову. У цьому випадку стає необхідним додатковий етап перекладу – з мови-посередника вхідного мови на мову-посередник вихідного мови. У такий спосіб досягається більш глибокий лінгвістичний аналіз і синтез. Переважна більшість сучасних комерційних СМП відноситься до систем трансферного типу.

Системи, ґрунтовані на використанні баз знань і мови-посередника, на сьогодні є системами майбутнього та знаходячись на етапі лабораторних досліджень і являючись, по суті, експериментальними. З точки зору користувача системи машинного перекладу можуть поділятися на три основних типи: інформативні, професійні та персональні.

Інформативні системи призначені для допомоги тим, кому потрібен доступ до інформації на іноземній мові і хто готовий користуватися «грубим», але досить зрозумілим перекладом. Такі системи, як правило, мають словники великого обсягу, але не спираються на новітні досягнення в лінгвістиці і програмуванні. Професійні дозволяють створювати нотатки для перекладу професійним перекладачам. Такі системи тепер використовуються все рідше, – як правило, при великому обсязі поточної перекладацької роботи, що виконується одночасно багатьма фахівцями в одній предметній області. Найчастіше в цих ситуаціях перекладачів постачають автоматичними словниками, тезаурусами з інтерактивним доступом або системами, які отримали назву «пам'ять перекладача». Персональні системи машинного перекладу зазвичай «працюють в діалозі» з користувачем і можуть давати задовільний переклад (якість якого все ж залежить від того, наскільки автор володіє вихідним мовою).

За мірою автоматизованості СМП можуть бути **автоматичними** і **автоматизованими**, у яких ряд функцій перекладу лишається за людиною. Детальніша класифікація за рівнем автоматизованості перекладу була запропонована у 1990 р. Ларрі Чайлдсом. Відповідно до цієї класифікації, виділяють наступні категорії МП: FAMT (Fully-automated machine translation) – повністю автоматизований машинний переклад; HAMT (Human-assisted machine translation) – машинний переклад за участю людини; MANT (Machineassisted human translation) – переклад, здійснюваний людиною з використанням комп'ютера. Повністю автоматизований переклад (FAMT) полягає в тому, що в систему автоматизованого перекладу вводиться текст однією мовою. Цей текст опрацьовується і виводиться текст іншою мовою. Реалізація такого виду автоматичного перекладу зустрічається з великою кількістю перешкод, найпоширенішими серед яких є дослівний переклад ідіоматичних виразів та неправильне визначення частин мови. Часто стверджують, що успіх

повністю автоматизованого перекладу потребує, в першу чергу, вирішення проблеми розуміння природної мови. Машинний переклад за участю людини (НАМТ) можливий тільки в умовах штучно обмеженої, як за словниковим запасом, так і за граматиною, мови. Не тільки комп'ютер перекладає весь документ, але й людина, яка усуває при цьому неясні, двозначні, складні конструкції (так звана «контрольована мова»). Це відбувається завдяки тому, що людина заздалегідь вкорочує довгі речення або коли людина повинна сама обрати значення слова. При НАМТ увесь процес перекладу здійснює людина, а комп'ютер автоматично перевіряє термінологію, здійснює пошук необхідних слів у словнику і порівнює отримане з іншими перекладами. Такі програми-перекладачі називають ТМпрограмами (від translation memory – пам'ять перекладу). Цю категорію програм застосовують професійні перекладачі. Основу ТМ-програм складають спеціалізовані словники, які відповідають тематиці тексту, що перекладається. При перекладі використовуються конструкції та значення слів і стійких словосполучень, обрані професійним перекладачем і занесені в словникові системи, а отриманий текст піддається редагуванню фахівцем-перекладачем. Словники і вже перекладені фрагменти текстів, що запам'ятовуються в ТМ-системі, можуть бути повторно в використанні у великих колективних проектах.

Лекція 2.

Алгоритм діяльності в процесі машинного перекладу

- 1) Рівні сприйняття мови.
- 2) Способи декодування.
- 3) Особливості машинного перекладу.

4) Етапи машинного перекладу.

Мовне спілкування за допомогою мови являє собою своєрідне **кодування** (говорити) і **декодування** (слухати) мовних сигналів (звуків, складів, слів, речень). У сприйнятті мови можна виділити два рівня, або дві сторони, цього двоєдиного процесу:

- 1) аналіз і синтез звуків
- 2) розуміння мови, або аналіз і синтез сигнальних, смислових характеристик мови.

При мовному спілкуванні відбувається кодування інформації комунікатором і її декодування реципієнтом. Кодування інформації, її мовне оформлення, передує усвідомленню змісту повідомлення. Отримання інформації супроводжується взаємообумовлених процесами декодування мовних сигналів і виявлення їхнього змісту.

Для переклада тексту необхідно знайти **спосіб декодування**, що використовує природну надмірність, у зв'язку із чим декодування повинне бути ймовірнісним. Завдання такого декодування полягає в тому, щоб, при даному повідомленні, знайти вхідне повідомлення, якому відповідає найбільша ймовірність. Для цього ж необхідно для будь-яких двох повідомлень уміти знаходити умовну ймовірність того, що перекладене повідомлення, пройшовши через канал із шумом, перетвориться у вхідне повідомлення. У цьому випадку потрібна **модель джерела** (модель мови) і **модель каналу** (модель перекладу). Модель мови дає оцінку ймовірності фразам перекладної мови, а модель перекладу оцінює ймовірність вихідної фрази за умови фрази перекладною мовою. Якщо нам потрібно перекласти фразу з української на англійську, то ми повинні знати, що саме зазвичай говорять англійською мовою і як англійські фрази спотворюються до стану української мови. Сам по собі переклад перетворюється в процес пошуку такої англійської фрази, яка максимізувала б добуток безумовної

ймовірності англійської фрази й імовірності української фрази (оригіналу) за умови даної англійської фрази.

Характерні риси та **особливості машинного перекладу**:

1. машинний переклад є своєрідним різновидом перетворення текстів однієї мови іншою;

2. при використанні машинного перекладу можна говорити про багатоступінчасту систему розкодування інформації, коли йдеться не тільки про переклад з однієї мови на іншу, але ще й переклад з мови комп'ютера на людську мову;

3. він можливий за наявності відповідного посередника, яким наразі є комп'ютер, спеціальна комп'ютерна програма, система машинного перекладу. Тому розвиток машинного перекладу упирається у розвиток інформатики, кібернетики та технологій;

4. якість машинного перекладу перебуває у прямій залежності від алгоритму, яким працює електронний перекладач, тобто від якості програми. Відповідно, чим краща якість програми, чим вона досконаліша, тим вищу якість перекладу можна отримати в результаті;

5. машинний переклад орієнтований насамперед на переклад текстів, хоча активно ведуться розробки з перекладу мовного потоку;

6. комп'ютерний переклад відрізняє висока швидкість обробки інформації, проте іноді це відбувається на шкоду якості;

7. використання машинного перекладу має на увазі повну або часткову відсутність перекладача-людини, тим самим людський фактор зведений практично нанівець, що часто відіграє важливу роль при перекладі конфіденційної інформації.

Алгоритм діяльності в процесі машинного перекладу включає такі фази:

- проводиться морфологічний та синтаксичний аналіз тексту, для кожного речення добираються різні варіанти перекладу, при цьому враховується вірогідність використання;
- на наступному етапі декодер за допомогою моделювання тієї чи іншої мови оцінює, як часто це речення вживається, і вибирає його.

Системи МП цілком можуть застосовуватися там, де використовується максимально стандартизована мова з простою граматиною й порівняно невеликим запасом слів. Досить успішним проектом системи МП вважається німецька програма «Meteo», що виконує переклад метеопрогнозів із французької мови на англійську й навпаки. Для полегшення роботи перекладачів і технічних письменників компанією Boeing у свій час був розроблений стандарт мови для написання технічної документації, який відомий як «Boeing English».

Машинний послідовний переклад. В галузі усного перекладу з'являються нові технології досягнень. Наприклад, «Skype Translator» пропонує користувачам спілкуватися виключно дистанційно, «Microsoft PowerPoint Presentation Translator» перетворює голосове повідомлення на текстове, а послідовний переклад через мобільний додаток від «Google Translate» нагадує спілкування людей із телефоном, ніж справжню розмову.

Машинний синхронний переклад. Ще у 1999 році японські науковці представили на конференції, присвяченій штучному інтелекту, експериментальну систему для англо-японського синхронного перекладу. Здійснюються спроби зробити технологічний аналог синхронного перекладача. Наприклад, розумні навушники в парі з мобільним додатком від «Waverly Labs». Якісний переклад досить простих фраз навіть для мовної пари англійська-іспанська ще не під силу електронному перекладачеві.

Програми RSI (розпізнавання голосу) уміють робити безліч корисного, причому з мінімальними помилками, якщо не йдеться про рідкісні акценти, сленг чи вузькоспеціальні теми. Програми використовуються у роботі при підготовці стенограм аудіозаписів для переклада, при диктуванні – від інтерв'ю до лекцій. Звичайно, тут важливі й зовнішні умови – наявність нормального чутливого мікрофона, мінімум сторонніх звукових перешкод і достатньо чітка (а ідеально – виразна) дикція спікера.

Лекція 3.

Особливості автоматизованого машинного перекладу

- 1) Із історії машинного перекладу.
- 2) Машинний переклад за участю людини.
- 3) Функції програм автоматизованого перекладу.
- 4) Обмеження та виклики автоматизованого перекладу текстів.

За останні 10 – 15 років характер роботи перекладача й вимоги до його підготовки істотно змінилися. У першу чергу зміни торкнулися письмового перекладу науково-технічної, офіційної й ділової документації. Сьогодні, як правило, уже недостатньо просто перекласти текст, користуючись комп'ютером як друкарською машинкою. Від перекладача очікують, що оформлення готового документа відповідатиме зовнішньому вигляду оригіналу настільки точно, наскільки це можливо, і при цьому має задовольняти прийняті у даній країні стандарти.

У наші дні, з огляду на бурхливий розвиток інформаційних технологій і неухильний ріст обсягу інформації, перекладацька діяльність стає все більш затребуваною. Сучасному майбутньому перекладачеві доводиться не тільки переглядати професійну та іншу літературу з метою витягу

інформації, але й обробляти її, тобто, практично, займатися перекладом. Як наслідок технічного прогресу – підвищення вимог і до підготовки майбутніх перекладачів, зокрема, до особливостей використання машинного перекладу в умовах їхньої професійної підготовки.

Вперше думку про можливість машинного перекладу висловив Чарльз Беббідж (1791-1871), який розробив у 1836-1848 рр. проект цифрової аналітичної машини – механічного прототипу електронних цифрових обчислювальних машин, що з'явилися через 100 років.

У 18 і 19 століттях було розроблено багато універсальних мов, у тому числі й Есперанто. Відомо, що ідея універсальної мови для машинного перекладу, ніяк не проявилася на початкових етапах розвитку цієї технології. Замість неї розглядалися тільки пари мов. Однак, протягом 1950-х і 60-х років, дослідники в Кембриджі на чолі з Маргарет Мастерман, у Ленінграді на чолі з Миколою Андрєєвим і в Мілані Сільвіо Чеккато почали роботу в цій галузі. У 1970-х і 1980-х роках були зроблені певні успіхи в цій галузі й був побудований ряд систем машинного перекладу. Фактично історія машинного перекладу починається з „Джорджтаунського експерименту“. У січні 1954 р. відбулася перша публічна демонстрація машинного перекладу з російської на англійську, здійсненого на машині IBM-701.

Інший напрямок робіт виник у Відділенні прикладної математики Математичного інституту Академії Наук СРСР за ініціативою А. А. Ляпунова. До робіт з машинного перекладу математичних текстів з французької мови на російську він залучив О. С. Кулагіну, аспірантку МІАН, своїх учениць Т. Д. Вентцель і Н. Н. Рікко. З кінця 1955 р. у цих роботах брала участь Т. Н. Молошная, яка пізніше приступила до самостійної роботи над алгоритмом англо-російського перекладу. Перші програми машинного перекладу, розроблені цим колективом, були реалізовані на машині „Стріла“.

Перше покоління систем машинного перекладу базувалося на алгоритмах послідовного перекладу „слово за словом“, „фраза за фразою“. Можливості таких систем визначалися доступними розмірами словників, прямо залежали від обсягу пам'яті комп'ютера. Переклад тексту здійснювався окремими пропозиціями, смислові зв'язки між ними ніяк не враховувалися. Такі системи називають системами прямого перекладу. На зміну їм з часом прийшли системи наступних поколінь, в яких переклад від мови до мови здійснювався на рівні синтаксичних структур. В алгоритмах перекладу використовувався набір операцій, що дозволяє шляхом аналізу перекладного пропозиції побудувати його синтаксичну структуру за правилами граматики мови вхідного пропозиції (так само, як навчають дітей мові в середній школі), а потім перетворити її в синтаксичну структуру вихідного пропозиції і синтезувати вихідну пропозицію, підставляючи потрібні слова зі словника. Такі системи називаються Т-системами (Т - від англійського слова „transfer - перетворення“). З 1960 р. почалася підготовка кадрів в галузі автоматичної переробки текстів.

Машинний переклад за участю людини розглядається як система, у якій комп'ютер виконує переклад як такий, але при цьому на різних стадіях виконання перекладу може взаємодіяти з людиною на стадії підготовки тексту до перекладу або на стадії виведення перекладу.

Першу стадію називають **передредагуванням**, другу - **постредагуванням**. Основною метою передредагування є виділення таких елементів як рідкісні словосполучення або ідіоми в процесі опрацювання тексту. Перекладач відповідним чином компенсує текст вхідною мовою.

Постредагування полягає в коригуванні перекладацького тексту, згенерованого системою машинного перекладу. Цю роботу виконує перекладач для того, щоб текст відповідав обумовленим термінам.

Інтерредагування – це втручання людини безпосередньо в процес перекладу, коли вирішуються складні питання.

Програми автоматизованого перекладу текстів є інструментами, що використовують штучний інтелект для перетворення тексту з однієї мови на іншу. Вони широко застосовуються в сучасному світі для спрощення комунікації між людьми, які розмовляють на різних мовах.

Програми автоматизованого перекладу текстів базуються на складних алгоритмах та машинному навчанні. Вони аналізують вихідний текст, розбивають його на фрази, слова та контекстуальні одиниці, і далі застосовують перетворення, щоб отримати перекладений варіант. Основна мета цих програм полягає в забезпеченні максимально точного та зрозумілого перекладу тексту.

Функції програм автоматизованого перекладу включають:

1. **Переклад тексту.** Це основна функція програм. Вони здатні перекладати текст з однієї мови на іншу швидко та ефективно. Користувачі можуть ввести вихідний текст або завантажити файл для перекладу.

2. **Розпізнавання мови.** Багато програм автоматизованого перекладу мають функцію визначення мови вхідного тексту автоматично. Це особливо корисно, коли користувач не знає, на якій мові написаний вихідний текст.

3. **Корекція граматичних помилок.** Деякі програми мають вбудовану функцію корекції граматичних помилок. Вони можуть виявляти та виправляти неправильну граматику або орфографію в перекладі. Це допомагає забезпечити більш точний результат.

4. **Синоніми та контекст.**

Деякі програми автоматизованого перекладу здатні використовувати синоніми та контекстуальний аналіз для поліпшення перекладу. Вони враховують контекст вхідного тексту, щоб вибрати найбільш підходящі переклади слів та виразів.

Переваги використання програм автоматизованого перекладу.

Використання програм автоматизованого перекладу текстів має численні переваги для користувачів. Ось декілька з них:

1. **Часова ефективність.** Завдяки автоматизованому перекладу текстів можна значно скоротити час, який був би витрачений на ручний переклад. Це особливо корисно при обробці великих обсягів тексту, таких як документи, веб-сторінки або електронні листи.

2. **Економія витрат.** Використання програм автоматизованого перекладу може допомогти зменшити витрати на професійні перекладачі або агентства перекладу. Вони є більш доступним і економічним рішенням, особливо для невеликих бізнесів або осіб, які потребують перекладу час від часу.

3. **Масштабованість.** Програми автоматизованого перекладу здатні обробляти великі обсяги тексту одночасно. Це означає, що їх можна використовувати для перекладу документів, веб-сторінок або навіть цілих сайтів. Це робить їх ідеальними для компаній та організацій з міжнародним присутністю.

Переваги використання програм автоматизованого перекладу текстів безсумнівні, однак, важливо враховувати також обмеження та виклики, з якими вони можуть стикатися.

Обмеження та виклики автоматизованого перекладу текстів. Хоча програми автоматизованого перекладу текстів є потужними інструментами, вони також мають деякі обмеження та виклики, які варто враховувати. Деякі з них включають:

1. **Полісемія та контекст.** Одним із найбільших викликів при автоматизованому перекладі є розрізнення полісемічних слів та правильне визначення контексту. Багатозначні слова можуть мати різні значення в різних контекстах, і точний переклад залежить від зрозуміння цього контексту.

2. **Граматичні та стилістичні помилки.** Програми автоматизованого перекладу не завжди здатні точно відтворити граматику та стиль оригінального тексту. Це особливо помітно при перекладі

складних фраз, ідіом або фразеологізмів. Важливо редагувати та коригувати переклад, щоб забезпечити правильну граматику та стиль.

3. Культурні та мовні відмінності. Кожна мова має свої унікальні культурні та мовні особливості. Це може призводити до труднощів у точному відтворенні нюансів та ідіоматичних виразів з однієї мови на іншу. Автоматизовані перекладачі можуть потребувати додаткової ручної настройки для врахування цих особливостей.

Незважаючи на ці виклики, програми автоматизованого перекладу текстів стають все більш точними та ефективними завдяки розвитку технологій машинного навчання та штучного інтелекту.

Перспективи розвитку програм автоматизованого перекладу. За останні роки програми автоматизованого перекладу текстів зазнали значного розвитку, але його потенціал ще не вичерпаний. Ось кілька перспектив розвитку цих програм:

1. Вдосконалення точності перекладу. Одним з основних напрямків розвитку є покращення точності перекладу. Застосування новітніх методів машинного навчання та глибокого навчання може призвести до поліпшення якості перекладу, зменшення граматичних помилок і кращого розпізнавання контексту.

2. Розширення мовних можливостей. Багатомовність є ще одним аспектом розвитку програм автоматизованого перекладу. Здатність перекладати між більшою кількістю мов може забезпечити більш широке застосування програм у різних країнах та культурах.

3. Інтеграція з іншими інструментами. Інтеграція програм автоматизованого перекладу з іншими інструментами, такими як текстові редактори, електронні таблиці або електронні поштові служби, може значно полегшити роботу з перекладом. Це може включати автоматичний переклад в реальному часі або підказки перекладу під час набору тексту.

4. **Покращення взаємодії з користувачем.** Розвиток програм автоматизованого перекладу також передбачає поліпшення взаємодії з користувачем. Інтуїтивний та зручний інтерфейс, підказки та рекомендації під час перекладу можуть зробити процес більш простим та ефективним.

Найкращі програми автоматизованого перекладу текстів. У світі існує багато програм автоматизованого перекладу текстів, але деякі з них відрізняються високою якістю та широкими можливостями. Ось кілька найкращих програм автоматизованого перекладу, які варто врахувати:

1. Google Переклад

Google Переклад – одна з найпопулярніших та найвикористовуваніших програм автоматизованого перекладу. Вона пропонує переклад текстів між багатьма мовами та має широкий функціонал, включаючи голосовий ввід та вивід, переклад в реальному часі та можливість редагування перекладу.

2. DeepL

DeepL є іншою потужною програмою автоматизованого перекладу. Вона відома своєю високою точністю та здатністю зрозуміти контекст перекладу. DeepL підтримує переклад текстів між багатьма мовами та надає можливість вибору різних стилів перекладу.

3. Microsoft Перекладач

Microsoft Перекладач – ще одна популярна програма автоматизованого перекладу. Вона пропонує переклад текстів між багатьма мовами та має інтеграцію з іншими продуктами Microsoft, такими як Word та Outlook. Microsoft Перекладач також має зручний мобільний додаток для перекладу на ходу.

4. PROMT Online

PROMT Online – це інша програма автоматизованого перекладу, яка забезпечує переклад текстів між різними мовами. PROMT Online пропонує різні варіанти перекладу, включаючи загальний, технічний та бізнес-

переклад. Вона також надає можливість перекладу веб-сторінок та документів.

Ці програми автоматизованого перекладу є дуже корисними для швидкого та загального перекладу текстів. Кожна з них має свої переваги та особливості, тому вибір програми залежить від ваших потреб та вимог. Перед використанням програм автоматизованого перекладу варто враховувати їх обмеження та пам'ятати, що вони не замінюють професійного перекладача. У складних випадках або для вимог високої якості перекладу краще звернутися до кваліфікованого фахівця.

Незалежно від того, яку програму ви виберете, важливо бути уважним та розуміти, що автоматизований переклад може містити неточності та помилки. Перекладені тексти повинні бути перевірені та скориговані, особливо якщо вони мають бути використані для професійних цілей.

З розвитком технологій машинного навчання та штучного інтелекту можна очікувати подальше поліпшення програм автоматизованого перекладу. Це дозволить забезпечити більш точні, природні та зрозумілі переклади для всіх користувачів.

Лекція 4.

Основні принципи декодування мовного сигналу

- 1) Проблема перекладності.
- 2) Трансформаційний підхід до перекладу.
- 3) Денотативний підхід до перекладу.
- 4) Принципи кодування мовних знаків.

Машинний переклад стикається з великою кількістю проблем, як специфічних (пов'язаних із суттю комп'ютерних систем і механізмів), так і загальних (перекладачі минулого порушували питання про перекладність). Такі знамениті особистості, як В. Гумбольдт або Г.В. Лейбніц, ставили під сумнів саму можливість адекватної передачі колориту однієї культури і національної своєрідності через мову іншої національності. І справді, проблема багатозначності під час машинного перекладу є однією з найскладніших прикладних задач, пов'язаних із лексичним значенням, навіть якщо брати до уваги, що існує й інший підхід – концепція універсальної перекладабельності, сформована на ідеях *трансформаційної граматики Н. Хомського*: "Усе, що виражено однією мовою, може - хай не без зусиль і не без втрат - бути виражено будь-якою іншою мовою... Сказане стосується не тільки раціональної інформації, але навіть "неперекладних" віршів, фразеологізмів і словесних каламбурів".

Принципи машинного перекладу можна розглядати в межах двох основних підходів, трансформаційного та денотативного.

Трансформаційний підхід розглядає переклад як перетворення об'єктів і структур однієї мови в об'єкти й структури іншої мови за певними правилами. У ході трансформації перетворюються об'єкти й структури різних мовних рівнів – морфологічного, лексичного, синтаксичного. Так, на лексичному рівні ми перетворюємо слова й словосполучення вихідної мови в слова й словосполучення мови перекладу. Тобто заміняємо одні структури іншими за певними правилами. Однак не можна забувати, що слова в складі словосполучень можуть перетворюватися інакше, ніж окремо взяті слова. Словосполучення – це вже маленький контекст, а контекст змінює значення слів і впливає на вибір еквівалента в мові перекладу. Таким чином, трансформації ми робимо під управлінням контексту. Наприклад, якщо трансформувати окремо взяте англійське слово "book", то можна з повною підставою замінити його головними словниковими еквівалентами –

іменником „книга” і низкою дієслів „замовляти”, „бронювати”, „резервувати”. Ці ж еквіваленти слова "book" залишаються і при перекладацьких трансформаціях більшості словосполучень із цим словом: "interesting book" - „цікава книга”, "book tickets" – „замовляти квитки” тощо. Однак, якщо ми трансформуємо, наприклад, словосполучення "book value", те одержимо зовсім інший український еквівалент „балансова вартість”, у якому немає еквівалентів окремо взятого слова "book".

Одна з проблем трансформаційного методу полягає в тому, щоб при перекладі за допомогою трансформацій відокремити зв’язані словосполучення від окремих слів, об’єднаних лише граматично, і зробити трансформацію відповідно до результатів такого поділу.

Надійного формального методу виділення зв’язаних словосполучень не існує, тобто, скажемо, для систем машинного перекладу, які здебільшого базуються на трансформаційному підході, тісніший зв’язок між словами "book" і "value" у словосполученні "book value" не помітний – для них таке словосполучення нічим не відрізняється, наприклад, від сполучення слів "book store" (книгарня). Людина ж виділяє словосполучення такого роду на основі складного аналізу змісту, а відповідний еквівалент зберігає в пам’яті або знаходить у словнику.

На синтаксичному рівні в процесі перекладу здійснюються трансформації синтаксичних конструкцій вихідної мови у відповідні конструкції мови перекладу.

Трансформаційний метод перекладу можна порівняти з розшифровкою зашифрованого тексту за допомогою „книги кодів”, роль якої виконує двомовний словник, і „зводу правил дешифрування”, викладених у граматичному довіднику.

Інший розповсюджений підхід до теоретичного тлумачення перекладацького процесу називається **денотативним**. Відповідно до цього

підходу, переклад здійснюється як трьохетапний процес, що складається з таких етапів:

- 1) сприйняття повідомлення вихідною мовою;
- 2) формування розумового образу (концепту) цього повідомлення;
- 3) інтерпретації цього образу засобами мови перекладу.

На відміну від трансформаційного, денотативний підхід не встановлює прямий зв'язок між словами й словосполученнями двох мов – переклад за денотативним механізмом припускає вільний вибір засобів мови перекладу для передачі змісту повідомлення вихідною мовою.

Назва цього методу походить від слова денотат, тобто фрагмент об'єктивної реальності, з яким співвідноситься як вихідне повідомлення, так і його переклад. Найбільш наочно цей підхід ілюструє переклад ідіом. У наведених нижче прикладах відсутність прямого зв'язку між вихідним текстом і його перекладом очевидна, вони зв'язані лише загальним змістом: "Can the leopard change his spots?" – „Горбатого могила виправить”. "There are many ways to skin a cat" – „Є багато засобів це зробити”. "Don't trouble trouble before trouble troubles you" – „Не буди лиха, поки лихо спить”.

Переклад за трансформаційним механізмом вимагає менше „розумових зусиль” і, як правило, перекладачі використовують його у своїй рутинній роботі, перекладаючи слово за словом, поки не натраплять на таке слово або на таку граматичну конструкцію, які змусять їх змінити порядок слів, перефразувати переклад або взагалі відмовитися від трансформацій і піти по шляху інтерпретації змісту оригіналу (тобто застосувати денотативний підхід).

При *синхронному перекладі* на інтерпретацію просто немає часу, тому синхроністи, як правило, переводять за трансформаційним механізмом, найчастіше жертвуючи стилістичною „гладкістю”.

При *усному послідовному перекладі*, коли потрібно запам'ятати і перекласти відразу кілька речень, природно, переважає денотативний підхід, тобто інтерпретація, і переклад рідко буває структурною копією оригіналу.

На вибір підходу, безумовно, впливає й жанр оригінального тексту – у загальному випадку при перекладі художньої літератури, особливо поезії, переважає денотативний підхід, тому що завдання такого перекладу не тільки й не стільки передати зміст, скільки створити адекватний образ, викликати в читача відповідні емоції й асоціації, а засоби для цього в різних мовах бувають різні.

При перекладі наукової і технічної літератури, навпаки, важливіше всього точно передати зміст і тут природно переважають трансформації.

Відомо, що мову можна представити у вигляді символного коду, за допомогою якого записуються розумові уявлення про речі реального світу (концепти). Наприклад, концепт будинку відповідає в українській мові ланцюжку символів БУДИНОК, а в англійській – послідовності символів BUILDING. Тоді, мабуть, можна сказати, що різні ланцюжки символів, тобто слова різних мов (наприклад, БУДИНОК і BUILDING) відповідають тому самому концепту (у нашій прикладі, концепту будинку). У такому випадку, якщо слово однієї мови, **A**, дорівнює концепту, **C**, і слово іншої мови, **B**, теж дорівнює цьому концепту, **C**, то за принципом транзитивності: Якщо **A=C** и **B=C**, то **A=B**. Інакше кажучи, елементи різних мов можна прирівняти на основі їхньої відповідності тому самому значенню.

На цьому принципі і заснована робота автомата, що перекладає, – слова різних мов прирівнюються один одному на основі спільності значення. Для створення автомата, що перекладає, відповідності між словами різних мов заносяться в програму комп'ютера і завдання такого примітивного автомата досить просте – для слів однієї мови він знаходить відповідності в іншій мові і так, слово за словом, перекладає текст. Однак, як вам теж повинне

бути відомо, спільне значення в різних мовах можуть мати не тільки окремі слова, але й словосполучення, тому найпростіший автомат, що перекладає, шукає відповідності не тільки для окремих слів, але і для словосполучень, виконуючи так званий послівно-пооборотний переклад.

Крім того, відповідність слів різних мов одне одному не є однозначною, тобто одному слову мови **A** може відповідати кілька слів мови **B** і навпаки. Отже, в автоматі, що перекладає, необхідно передбачити програму вибору правильного еквіваленту.

Такі програми звичайно ґрунтуються на двох принципах:

- (1) на принципі вибору еквіваленту за синтаксичною моделлю вхідного тексту, наприклад, автомат може розрізнити закладені в його словник еквіваленти дієслова *to book* і іменника *a/the book* (відповідно, „резервувати” і „книга”);
- (2) на складнішому принципі вибору еквіваленту за семантичною моделлю. За різними семантичними моделями автомат, наприклад, може розрізняти такі еквіваленти слова *order*, як „порядок” і „замовлення”.

Обидві моделі звичайно застосовують у комплексі. У деяких складніших системах на додаток до цих двох принципів вибору еквіваленту застосовують також і принцип вибору на основі екстралінгвістичної (фонової) інформації. Моделі для вибору еквівалентів за цим принципом відносять до розряду моделей штучного інтелекту.

Лекція 5.

Багатозначність як проблема автоматизованного перекладу

- 1) Проблема лексичної багатозначності.
- 2) Причина помилок в перекладі.
- 3) Проблема вирішення лексичної багатозначності.

4) Принципи створення систем машинного перекладу текстів.

Проблема вирішення *лексичної багатозначності* є однією з найскладніших прикладних задач, пов'язаних з лексичним значенням. Завдання автоматичного (рідше напіваавтоматичного) вирішення лексичної багатозначності була вперше сформульована в рамках напрямку науки і технології, пов'язаного зі створенням систем машинного перекладу. Надалі проблема вирішення лексичної багатозначності стала однією з ключових не тільки при створенні систем машинного перекладу, а й систем обробки текстів природною мові інших призначень (пошук, класифікація).

Полісемія (від грец. Polysemos – багатозначний) – наявність у одиниці мови більш одного значення – двох або декількох. Часто, коли говорять про полісемію, мають на увазі багатозначність слів як одиниць лексики. **Лексична полісемія** – здатність одного слова служити для позначення різних предметів і явищ дійсності. Багатозначність, багатоваріантність, тобто наявність у слова (одиниці мови, терміна) двох і більше значень, історично обумовлена.

Омонімія, і полісемія проявляються в тому, що одна звукова одиниця має кілька різних значень. Разом з тим відмінність полісемії і омонімії полягає в наступному: в разі полісемії всі значення даного слова утворюють певну систему, відносини між усіма ними усвідомлюються мовцями. При **омонімії** значення слів не пов'язані між собою. Якщо розглядати різновиди полісемії, то виділяють граматичну та лексичну полісемію, а також полісемію морфем. Лексична полісемія – здатність одного слова служити для позначення різних предметів і явищ дійсності, є найпоширенішим і найбільш уживаним видом. Наприклад, іменник «ball» має більше 10 лексичних значень серед яких: 1) м'яч; 2) сфера; 3) ядро, знаряд; 4) клубок шерсті тощо. Те, в якому із значень виступає слово, визначається його сполучністю з іншими словами: "inaugural ball", "crystal ball", "soccer ball".

Реалізацію того чи іншого значення полісеми здійснює контекст, ситуація або загальна тематика тексту. З іншого боку, за певних умов контекст може створювати так звану семантичну дифузність, тобто сумісність окремих лексичних значень, коли їхнє розмежування не здійснюється, бо не є необхідним. Деякі значення проявляються лише у поєднанні з означальним словом (inaugural ball); в окремих випадках багатозначне слово може набувати фразеологічного відтінку, наприклад, «drop the ball». Найбільш багатозначними словами в англійській мові є такі слова, як make, get, run або set. Деякі джерела виділяють до кількох сотень значень. Так, наприклад, в англо-російському словнику Мюллера set як дієслово має 26 значень, іменника – 24, прикметника – 7.

Основною *причиною помилок* стають неправильно підібрані синонімічні варіанти перекладу багатозначних слів, наприклад:

- Meal – їжа; стіл; страва; сніданок; обід; трапеза; борошно; частування; корм.
- To top – перевершувати; покривати куполом; перевищувати; досягати будь-якої величини; піднятися на вершину; завершувати.
- Wave – махання; хвиля, хвилястість; коливання; нерівність; підйом.
- Pipe – трубка; люлька; сопілка; дудка; свисток; трубопровід; винна бочка; труба.
- Unrest – занепокоєння; хвилювання; заворушення; хаос.
- Scramble – сутичка; боротьба; бійка.

Проблема вирішення лексичної багатозначності системами машинного перекладу полягає у взаємодії загальноновживаних та предметних словників. Так, наприклад, практично в кожному науковому тексті можуть бути описи, що виходять за межі цієї конкретної предметної області. Тобто у медичному тексті, що володіє власною лексикою, можливий опис нового технічного пристрою, і в цьому описі може з'явитися велика кількість технічних термінів, що не входять до традиційної лексики медичної сфери.

Одна з найсуттєвіших проблем систем машинного перекладу – використання найбільш вживаного значення слова, що не завжди є доцільним. Наприклад, іменник *runners* у контексті речення *The sled was without runners* потребує аналізу з використанням тематичних словників.

У прикладних задачах комп'ютерної лінгвістики не робиться відмінності між омонімічними і полісемічними значеннями слова. Це пов'язано з тим, що в переважній більшості прикладних задач важлива не стільки етимологія слова, скільки його семантика. Розпізнавання і поділ груп омонімічних значень також входить в завдання вирішення лексичної багатозначності, так як іноді може виявитися корисним з практичної точки зору.

Проблема багатозначності вважається вирішеною, якщо для слова вибрано його регулярне значення або якщо знайдений синонімічний еквівалент у вигляді регулярного значення для метафоричного використання. Відомо, що при вирішенні багатозначності існує ряд самостійних завдань. Зокрема, можна виділити найбільш великі, «класичні» завдання: 1) Завдання приписування відомого значення відомій лексемі. 2) Завдання приписування відомого значення новій лексемі. 3) Завдання виявлення нового значення для відомої лексеми. 4) Завдання виявлення нового значення для нової лексеми. Розрізняють два основні класи механізмів вирішення багатозначності:

1 клас. Це механізми автоматичні, які передбачають повністю комп'ютерне рішення цього завдання.

2 клас. Це механізми інтерактивні (діалогові, напіваавтоматичні), які передбачають спільне рішення задачі людиною і комп'ютером, і зводяться до того, що комп'ютер надає користувачу набір альтернатив, з якого він повинен вибрати один варіант. Одним з автоматичних методів розв'язання багатозначності є фільтри, тобто методи, які не виявляють точного значення, але в явному вигляді накладають обмеження на їх спектр.

Граматична полісемія – збіг різних граматичних форм однієї лексеми. Реалізацію того чи іншого значення слова здійснює контекст або ситуація, загальна тематика мови. Точно так же, як контекст обумовлює конкретне значення багатозначного слова, в певних умовах він може створювати семантичну диффузність, тобто сумісність окремих лексичних значень, коли їх розмежування не проводиться (і не представляється необхідним). Деякі значення проявляються тільки в поєднанні з визначальним словом; в деяких поєднаннях значення багатозначного слова представлено як фразеологічно пов'язане. Наприклад, у реченні *The dog is playing outside*. (Ця собака грається на дворі) артикль виконує уточнюючу функцію, а в реченні *The dog is a domestic animal* (Собака – домашня тварина) – узагальнюючу. Поряд з лексичною та граматичною полісемією існує полісемія морфем. Вона виникає завдяки здібності морфем мати декілька значень, між якими існує семантичний зв'язок. Наприклад, у слові *being* (життя, існування), закінчення - *ing* може бути показником процесу або приналежності слова до класу іменників.

Синтаксичні трансформації в машинному перекладі.

Дієслівно-іменникові трансформації – центральне питання формування структури перекладного висловлювання. Новий зміст проблемі мовних трансформацій надають сучасні реалії: необхідність проектувати і розвивати навчальні компоненти систем машинного перекладу й редагування текстових знань на основі вже існуючих і знову створюваних корпусів паралельних текстів. На сучасному етапі лінгвістичних досліджень і розробок необхідно синергетичне поєднання функціонального і рівневого підходів.

Функціональний підхід інтегрує мовні засоби (синтаксичні, лексичні, словотворчі і словозмінні), що належать різним рівням мови, на основі їх функціонально-семантичних характеристик. Під трансформаціями розуміються, перш за все, перетворення предикаторов в імена і імен в

предикатори: бігти – біг, вчитель – вчителювати, при цьому зберігається часткова тотожність форма – корінь або основа слова і певна тотожність семантики. Трансформації постійно виступають як одне з двох головних засобів – поряд з перифразами – створення висловлювань.

Переклад фразеологічних поєднань в системах машинного перекладу. Проблема машинного перекладу ідіом полягає в тому, що не завжди вдається дати точний переклад, керуючись звичайними правилами. Однак слід взяти до уваги, що ідіоми повинні бути виявлені на початковому етапі щоб уникнути їх втрати, і оброблятися вони повинні як одне слово. Вирішення ідіоматичності є однією з формальних операцій, що забезпечують аналіз і синтез в системах машинного перекладу, і виробляються або за допомогою стандартних граматичних і лексичних програм аналізу тексту, що працюють спільно з автоматичним словником, або шляхом прямого співвіднесення вхідного і вихідного сегментів. У другому випадку і вхідний, і вихідний сегменти розглядаються як неподільні звороти. Один або кілька вихідних зворотів або словоформ, поставлених у відповідність кожному вхідному звороту, складають автоматичний словник зворотів. В результаті аналізу ідіоматичним виразами приписується певний цифровий еквівалент, і вони виключаються з подальшого граматичного аналізу.

Складанню алгоритму пошуку і перекладу зворотів в тексті передують лінгвістичне дослідження їх дистрибуції. Слід з'ясувати за допомогою синтаксичного аналізу, чи є зворот цілісним і чи включає він у себе змінювані форми. Зворот вважається цілісним, якщо він має постійний склад і між його елементами можна вставити інші одиниці. Якщо зворот не є цілісним, слід враховувати це при складанні алгоритму.

Таким чином, для обробки зворотів, розірваних іншими членами речення, необхідні дані синтаксичного аналізу. При створенні систем

машинного перекладу для текстів, що містять ідіоматичні вирази, необхідно керуватися наступними *принципами*:

1. Основними одиницями мови і мовлення, які слід включати в машинний словник, повинні бути фразеологічні одиниці (зокрема, ідіоматичні вирази). Окремі слова також можуть включатися в словник, але вони повинні використовуватися тільки в тих випадках, коли не вдається здійснити переклад, спираючись тільки на фразеологічні одиниці.

2. Поряд з ідіоматичними виразами, що складаються з безперервних послідовностей слів, в системах машинного перекладу слід використовувати і так звані "мовні моделі" - фразеологічні одиниці з "порожніми місцями", які можуть заповнюватися різними словами і словосполученнями, породжуючи осмислені відрізки мовлення.

3. Реальні тексти, незалежно від їх приналежності до тієї чи іншої тематичної галузі, зазвичай бувають політематичні, якщо вони мають чималий обсяг. І відрізняються вони один від одного не стільки словниковим складом, скільки розподілами ймовірностей появи в них різних слів з загальнонаціонального словникового фонду. Тому машинний словник, призначений для перекладу текстів навіть тільки з однієї тематичної галузі, повинен бути політематичним, а для перекладу текстів з різних предметних галузей – тим більше.

4. Необхідні машинні словники великого обсягу. Такі словники повинні створюватися на основі автоматизованої обробки двомовних текстів, які є перекладами один одного, і в процесі функціонування систем перекладу.

5. Поряд з основним політематичним словником великого обсягу, в системах фразеологічного машинного перекладу доцільно використовувати також набір невеликих за обсягом додаткових тематичних словників. Додаткові словники повинні містити тільки ту інформацію, яка відсутня в основному словнику (наприклад, інформацію про пріоритетні перекладних

еквівалентах словосполучень і слів для різних предметних галузей, якщо ці еквіваленті не збігаються з пріоритетними перекладними еквівалентами основного словника).

6. Поряд з перекладом текстів в автоматичному режимі, в системах фразеологічного машинного перекладу доцільно передбачити і інтерактивний режим їх роботи. У цьому режимі користувач повинен мати можливість втручатися в процес перекладу і налаштовувати додаткові машинні словники на тематику перекладаються текстів.

Лекція 6.

Помилки машинного перекладу писемних текстів та способи їх виправлення

- 1) Типи адекватності вперекладі.
- 2) Поняттєва структура змісту природно-мовного тексту.
- 3) Види перекладацьких помилок.
- 4) Шляхи уникнення помилок в машинному перекладі текстів.

Текст має бути зрозумілим і доступним читачеві, тобто бути адекватним.

Жанрово-тематична адекватність – отримання найзагальнішої інформації про першоджерело, а саме про його жанрову та тематичну співвіднесеність.

Семантико-синтаксична адекватність – еквівалентність термінологічних складових першоджерела і перекладеного текста.

Інформативна адекватність – відтворення основної інформації про зміст повідомлення, що міститься у першоджерелі.

Поняттєва структура змісту природно-мовного тексту є ієрархічною структурою, на верхньому рівні якої знаходяться найбільш загальні поняття і відношення між ними, кожний нижній рівень представляється поняттями і відношеннями, які конкретизують відповідні поняття і відношення найближчого вищого рівня. Верхній рівень відповідає найбільш загальному опису змісту тексту, нижчі її рівні відповідають рівням конкретизації цього опису. Кожне поняття і відношення в понятійній структурі тексту супроводжується характеристиками, які відображають їхні властивості (понять і відношень), модальності та інші аспекти; лінгвістичною інформацією, яка характеризує мовні засоби їх відображення у вхідному тексті; семантичною інформацією, яка відбиває їх роль та інші характеристики (наприклад, об'єкт, суб'єкт, тип відношення, напрямок дії тощо).

Необхідними умовами збереження форми та смислу текстових фрагментів є те, що, по-перше, машинна система розпізнає прийоми створення стилістичної одиниці, по-друге, відшукує та підбирає подібні стилістичні елементи для відтворення первинного семантичного значення у друготворі.

Проблема **оцінки якості перекладів** є надзвичайно актуальною. При визначенні природи **перекладацьких помилок** дослідники розглядають такі аспекти:

- 1) розмежування перекладацьких та мовних помилок;
- 2) створення переліку можливих перекладацьких помилок;
- 3) відносність природи перекладацьких помилок;
- 4) необхідність оцінки якості не лише на лінгвістичному, але й на прагматичному рівні.

Перекладацькі помилки впливають на передачу змісту тексту оригіналу в тексті-перекладі, порушують функції перекладу, зв'язність

тексту, лінгвістичні та культурно-специфічні умовності. Так, У. Беннет виокремлює чотири основних види перекладацьких помилок:

- 1) прагматичні, які виникають в результаті неадекватного вирішення прагматичних проблем перекладу;
- 2) культурні, які є результатом невірного відтворення специфічних культурних умовностей;
- 3) лінгвістичні, які виникають в результаті неадекватного перекладу мовних структур;
- 4) текстові, які пов'язані з проблемами перекладу певного тексту.

Американська асоціація перекладачів розробила наступні **критерії оцінки якості перекладів** претендентів на отримання сертифікату цієї організації. При оцінюванні якості перекладу помилки, які впливають на значення тексту оригіналу (тобто перекладацькі помилки), отримують різну кількість балів в залежності від ступеню викривлення значення та наслідків, до яких це викривлення призводить. Так, якщо помилка лише незначним чином впливає на розуміння, вона оцінюється в 1 бал, помилка з мінімальним втручанням в зміст оригіналу отримує 2 бали, якщо порушення змісту носить обмежений характер – 4 бали. Нарешті, якщо при всій серйозності помилки текст в цілому все ще є придатним до використання, така помилка отримує 8 балів, якщо ж ні – то 16 балів. Механічні помилки (тобто такі, які можна виявити без звертання до тексту оригіналу і які не впливають на зміст) не відносяться до дуже серйозних і оцінюються в 1 бал, у 2 бали – якщо є очевидною, і у 4 бали, якщо вимагаються певні зусилля для розуміння того, що мається на увазі під помилково вжитим словом. Наразі існує кілька спеціальних програм, які використовуються для оцінки якості перекладу. Одна з них, SAE J2450 Translation Quality Metric, розроблена Товариством інженерів-автомобілістів. Ця програма здійснює оцінку якості перекладу відповідно

до кількості та серйозності помилок в тексті-перекладі порівняно з оригіналом.

Помилки поділяються на сім типів:

- 1) некоректне вживання терміну;
- 2) некоректне значення лексичної одиниці;
- 3) пропуск;
- 4) помилки у структурі речення чи узгодженні слів у реченні;
- 5) некоректне написання слова;
- 6) пунктуаційні помилки;
- 7) інші помилки.

Помилки стилю не беруться до уваги, так само, як і помилки форматування. Типи помилок в свою чергу поділяються на підкатегорії в залежності від ступеня їхньої серйозності (серйозні або незначні). Кожному типу помилки присвоюється певна кількість балів, і чим меншим є бал, тим менше помилок містить переклад і, відповідно, тим кращою є його якість.

Оцінка якості перекладу, яка здійснюється за допомогою вищезгаданої програми, відображає ступінь відносної еквівалентності лінгвістичних елементів тексту-оригіналу та тексту-перекладу, при цьому імпліцитно припускається, що значення є функціонально еквівалентним як в контексті вживання тексту-оригіналу, так і тексту-перекладу.

Крім того, можна розділити на помилки сприйняття, помилки породження тексту і інтерферентними помилки. Помилки сприйняття найчастіше виникають в результаті помилкового розуміння синтаксичної структури оригіналу.

Приклади помилок машинного перекладу.

1. **Неправильне вирішення проблем лексичної полісемії:** англ. *Revisedwing has a 25% larger area, the spread is increased by 4.88 m, which increasedfuel capacity by 30%.*

Переглянуте крило має на 25% більшу площу, розповсюдження збільшено на 4,88 м, що збільшило ємність пального на 30%.

Адекватний варіант перекладу: Перероблене крило має на 25% більшу площу, розмах збільшений на 4,88 м, дозволило збільшити запас палива на 30%.

2. Неправильне вирішення проблеми лексично-граматичної омонімії. *It currently mines iron ore from its one open-pit quartzite field through a process of drilling and blasting, and by the removal of overburden to external dumps.*

В даний час вона видобуває залізну руду з одного відкритого кварцитового родовища в процесі буріння та вибухових робіт, а також шляхом вивезення розкривних порід на зовнішні звалища

3. Дослівний переклад словосполучень. *Under the Charter, the Security Council has primary responsibility for the maintenance of international peace and security.*

Відповідно до Хартії Рада Безпеки несе головну відповідальність за підтримку міжнародного миру та безпеки.

Для отримання кращої якості машинного перекладу засобами онлайн-сервісів та комп'ютерних програм перекладачів під час перекладу з англійської мови і навпаки необхідно дотримуватись попереднього редагування текстової інформації:

1. Уникнення друкарських та орфографічних помилок. Комп'ютерний перекладач не може виправити помилки та розпізнати неправильно надруковані слова. Тут стають у нагоді спеціальні програмні засоби для перевірки правопису, так звані Spellcheckers. Останні встановлюються на персональному комп'ютері, використовуються як безкоштовний онлайн-сервіс чи плагін до текстових редакторів. Найчастіше програми для перевірки правопису дозволяють перевіряти граматику. Як і онлайн словники, онлайн-перевірка правопису дозволяє

користувачеві друкувати слова, рядки, параграфи, перевіряючи автоматично правопис українською та іншими мовами.

2. **Дотримання правил пунктуації.** Пропущені або, навпаки, зайві розділові знаки перешкоджають електронному перекладачеві правильно зрозуміти синтаксичну структуру речення.

3. **Дотримання регістру.** Мала літера в слові легко перетворюється на прописну (велику) літеру, наприклад, на початку речення або в заголовку. Зазначене вище правило враховується під час розробки систем МП. І, навпаки, прописна літера стає малою дуже рідко.

4. **Уникнення зайвих слів.** Зазначене вище правило стосується видалення слів, які не розширюють змістового наповнення речення.

5. **Повторення іменників замість займенників.** Це правило сприяє адекватному «сприйняттю» змісту речень програмами перекладачами.

6. **Не використовувати аббревіатури.** Неправильний переклад аббревіатури вважається лише частиною проблеми. Зазначене полягає в тому, що навіть одне неправильно перекладене слово перешкоджає електронному перекладачеві правильно проаналізувати синтаксичну структуру речення (тобто, аббревіатури беруть участь у синтаксичних зв'язках поряд із звичайними словами).

7. **Уникнення сленгових виразів.** Зазначене вище правило стосується злочинного сленгу, хоча злочинці можуть застосовувати системи МП. У неформальній комунікації навіть законослухняні носії мови найчастіше слова, вирази та конструкції, що не належать до літературної мови.

8. **Уникнення кліше, ідіом та розмовних фраз.** Засоби МП, зазвичай, неправильно (найчастіше, буквально) передають їхній зміст. Тоді, у цьому випадку, речення мовою перекладу не матиме сенсу.

9. **Використання простих синтаксичних конструкцій** з прямим порядком слів.

10. Уникнення складних речень та однорідних частин речень.

Зазвичай, кожне речення слугує для передачі однієї логічної думки. Це правило, яке однаково застосовується для усіх мов, є найефективнішим. Треба уникати сполучників («and», «but», «which», etc.) та випадків уживання більш одного підрядного простого речення у межах складного. Рекомендований обсяг речення - не більш 25 слів. Кожне речення повинно бути граматично завершеним.

11. Вживання інфінітивних конструкцій замість герундіальних.

12. Використання активного стану дієслів замість пасивного.

Активний стан вважається прямим стилем написання, який допомагає уникнути розмитості та неоднозначності. Програмам МП дуже складно перекладати розмиті фрази або слова з подвійним значенням.

Постредагування передбачає удосконалення тексту, що отримано у процесі МПГ, редактором, який, зазвичай, проходить спеціальну підготовку та має досвід роботи з текстовими документами такого типу. Кількість часу та праці, які необхідні для постредагування, є одним із ключових факторів, які враховуються під час оцінки економічної ефективності МП. Літературні, рекламні та інші тексти, зазвичай, не рекомендовані для МП, не підлягають постредагуванню. Для того, щоб забезпечити належну якість тексту, наближену до людського перекладу, редакторам потрібно переписати текст знову. У цьому випадку переваги МП практично дорівнюють нулю.

Процес постредагування тексту МП включає такі етапи:

1. Перечитування речення чи фрагменту тексту МП.
2. Порівняння вихідного тексту МПз оригінальним текстом.
3. Оцінювання якості кожному фрагменту тексту перекладу на основі стандартизованих інструкцій.
4. Прийняття рішення щодо достовірності перекладу того чи іншого фрагмента.
5. Удосконалення або повторний переклад тексту.

Машинний переклад з постредагуванням вважається доцільним під час роботи з великими за обсягом та подібними за формою текстами.

Лекція 7.

Розпізнавання письмових текстів. Членування мовного потоку

- 1) Лінгвістичні бази даних.
- 2) Структурна організація корпусів.
- 3) Поняття «вирівнювання тексту».
- 4) Членування та об'єднання речень при перекладі.

Лінгвістична база даних для системи машинного перекладу включає в себе накопичені лінгвістичні дані, з текстами, картотеками, словниками, граматиками і іншими лінгвістичними джерелами. Типовий склад *лінгвістичної бази даних* можна обмежити наступними компонентами:

1) *Лексико-граматичний класифікатор* властивостей вихідного мови і мови перекладу (система морфологічного кодування). при аналізі вихідного

тексту кожне слово в ньому має отримати відповідні морфологічні характеристики: ознака частини мови, рід, відмінок, число і ін. Система кодування повинна бути єдиною для конкретної системи машинного перекладу.

2) *Базовий двомовний морфологічний словник*. У цьому словнику встановлюється послівна відповідність кожної словоформи вихідного мови словоформам мови перекладу.

3) *Словник скорочень і аббревіатур*. Словник використовується на етапі розбиття вихідного тексту на слова і речення. Скорочення і

аббревіатури повинні бути по можливості розшифровані, так як вони можуть бути членами речення, отже, їх необхідно враховувати при синтаксичному і семантичному аналізі.

4) **Словник ідіом.** Даний словник застосовується до синтаксичного аналізу, оскільки дуже часто ідіома є одним членом речення і розглядається як єдине ціле; при перекладі ідіома мовою оригіналу може відповідати одному слову мови перекладу.

5) **Термінологічні словники** за предметними галузями. Додаткові словники підключаються при необхідності перекладу спеціалізованих текстів.

6) **Синтаксичний словник.** В даному словнику повинна міститися інформація про синтаксичну сполучуваність членів речення як в мові оригіналу, так і в мові перекладу, а також синтаксичні відповідності, необхідні при перекладі.

7) **Семантичний словник** (тезаурус, онтологія). Даний компонент містить інформацію про семантичну сполучуваність лексем, про лексико-семантичні поля, які застосовуються на етапі побудови семантичної граfi речення.

8) **Корпус паралельних текстів.** Корпус містить тексти мовою оригіналу і їх переклади на іншу мову. При знаходженні речення або його фрагмента в корпусі паралельних текстів в текст перекладу вставляється його відповідник мовою перекладу. На використанні корпусу текстів побудована технологія пам'яті перекладів.

Паралельний корпус – це електронний анотований корпус, до якого крім оригінальних текстів залучено їхні переклади тією чи іншою мовою, які вирівняні за реченнями або абзацами.

Для паралельних корпусів виділяється ряд підтипів: тексти на мові А і їх переклади на мову В; тексти на мовах А і В і їх переклади відповідно на

мови В і А; тільки перекладні тексти на мовах А, В, С, Х, якщо оригінальні тексти були написані на мові D.

Крім того, до паралельних корпусів можна віднести діяхронічні корпуси, які складаються з текстів на більш ранній формі мови і їх перекладів на сучасну мову, транскрипційні корпуси текстів, що включають тексти літературною мовою, прочитані носіями різних її діалектів. В якості підтипів можна виділити «галасливі» паралельні корпуси; з пропусками в перекладі, без точної відповідності між оригіналом і перекладом), «дзеркальні» паралельні корпуси, що складаються з текстів на мовах А і В і перекладів цих текстів відповідно на мови В і А.

Важливість паралельних корпусів обумовлена тим, що вони дозволяють об'єктивно встановити, як перекладачі на практиці долають труднощі, і використовувати ці дані для розробки моделей для початківців перекладачів, що відповідають реальності. Вони також відіграють важливу роль в дослідженні перекладацької норми в специфічних соціокультурних та історичних контекстах. Це робить багатомовні корпуси текстів у багатьох відношеннях привабливими для перекладачів-практиків.

Паралельні корпуси залишаються незамінним джерелом даних для проведення досліджень в галузі прикладної лінгвістики (апробація систем автоматизованого перекладу, заповнення систем перекладацької пам'яті, розробка систем автоматичного пошуку перекладних еквівалентів і т.п.), так і для контрастивних і перекладознавчих досліджень (порівняння структури вихідного тексту і перекладу, визначення ступеня інформаційних втрат при перекладі, вивчення різних перекладацьких стратегій і т. п.)

При складанні паралельних корпусів, на відміну від одномовних і порівняльних корпусів текстів, слід враховувати фактор міжкультурних зв'язків. Безліч текстів вихідної мови складають лише ті тексти, які були переведені на іншу мову, і, якщо міжкультурні зв'язки повністю відсутні, отримання паралельного корпусу неможливо. Чим слабкіше

міжнаціональні та культурні зв'язки, тим менше перекладів виконується і тим більш проблематичним є складання повноцінного паралельного корпусу. Тексти на вихідній мові, хоча і є первинними, відбираються з урахуванням мови перекладу. Структура субкорпусу вихідного мови визначається наявністю або відсутністю перекладів на мову перекладу, а також тим, якого роду тексти перекладаються.

При складанні паралельних корпусів можуть використовуватися різноманітні мовні ресурси: спеціальні тексти, тексти ЗМІ, наукові тексти, художні тексти, тобто паралельний корпус повинен мати властивість *репрезентативності*.

Структурна організація корпусу може бути найрізноманітніша, в залежності від прагматичних цілей його автора:

- у вигляді традиційного тексту з посиланням на переклад;
- в табличній "дзеркальній" формі, що більш зручно для сприйняття і порівняння;
- у вигляді бази даних (структура, застосовна тільки при автоматичній обробці).

Важливим поняттям є *вирівнювання тексту*. Вирівнювання паралельного тексту – це ідентифікація відповідних один одному речень в обох половинах паралельного тексту. Вирівнювання паралельного корпусу на рівні речень є необхідною передумовою для різних аспектів лінгвістичних досліджень. В процесі перекладу речення можуть розділятися, зливатися, віддалятися, вставлятися або змінювати послідовність. У зв'язку з цим вирівнювання часто стає складним завданням.

Паралельні корпуси текстів-зразків (у вигляді бази даних) особливо корисні в тому випадку, коли перекладач працює зі строго нормованими текстами, жанрово-стилістичний і стилістичне оформлення таких текстів практично не допускає варіювання, відступу від певних соціокультурних

норм. Це тексти ділової листування, тексти-рецепти, тексти-прогнози погоди, тексти-контракти та т.д. Тексти різних стилів розрізняються як словником лексичних одиниць, що вживаються в певних текстах, так і граматичними і синтаксичними структурами речень.

Необхідною функцією сучасної системи автоматизованого перекладу є застосування *перекладацької пам'яті* — двомовної бази даних перекладених вихідних текстів і їхніх відповідників цільовою мовою, яка пропонує схожі наявні сегменти під час перекладу нових текстів.

Перекладацька пам'ять ділить вихідний текст на одиниці, які називаються *сегменти*. Речення або аналогічна одиниця (наприклад, заголовок, елемент списку, посилання тощо) з вихідного тексту вважається окремим сегментом. Під час роботи перекладача над документом програма з перекладацькою пам'яттю по черзі відображає кожен вихідний сегмент та пропонує повторно використати попередні переклади, якщо в базі даних наявний такий самий або схожий вихідний сегмент. Якщо його немає, перекладач може самостійно ввести власний варіант перекладу для нового сегмента. Після завершення роботи над сегментом переклад зберігається до бази даних, і програма переходить до наступного сегмента. Перекладацька пам'ять переважно є простою базою даних, що містить поле із сегментом вихідною мовою, поле із сегментом цільовою мовою, а також поле з метаданими, як-от дата створення, дата останнього використання, назва перекладача / проекту / файлу тощо.

Іншим варіантом створення перекладацької пам'яті є співставлення існуючих перекладених документів, замість їх поступового утворення й розширення під час роботи. Такі можливості спричинили виникнення нових перекладацьких стратегій, які включають нелінійну обробку та переклад текстів, що задовольняє вимоги до глобального поширення вмісту. Новоутворені тексти стають нелінійними, оскільки вони «запозичені» із вже існуючих.

Членування та об'єднання речень при перекладі. Необхідно розрізняти зовнішнє і внутрішнє розчленування. При **зовнішньому розчленуванні** одне речення оригіналу перетворюється у перекладі в два або декілька речень. При внутрішньому розчленуванні просте речення оригіналу замінюється складним. При перекладі з англійської мови набагато частіше доводиться вдаватися до розчленування, ніж до об'єднання речень. Це пояснюється централізованістю структури англійських речень. Внутрішнє розчленування речень буває необхідним при перекладі синтаксичних комплексів (інфінітивних, дієприкметникових, герундіальних), які перекладаються складними реченнями:

The danger of flooding is believed to be over. Є підстави вважати, що небезпека повені минула.

They saw the whole town shattered to a heap of ruins. Вони бачили, як ціле місто перетворилося у руїни.

Your having been ill makes no difference. Те, що ви були хворі, нічого не міняє.

Що ж стосується **внутрішнього розчленування**, то його необхідність нерідко диктується слабкістю смислових зв'язків між частинами англійського речення, що пояснюється тенденцією англійського синтаксису до об'єднання в одному реченні не зовсім однорідних думок. Деколи збереження структури речення ускладнюється різноплановістю думок, які містяться у ньому, особливо, коли в одному реченні об'єднані повідомлення, котрі відносяться до різного часу:

And here at times, when his mother's father's financial difficulties were greates, they would be found thinking, or as Asa Griffith was wont helplessly to say at times, praying their way out, a rather ineffectual way, as Clyde began to think later. І тут же в періоди найбільших фінансових труднощів батько й мати сиділи і роздумували, або як деколи безпомічно говорила Ейса

Гріфітс, молили бога вказати їм вихід із положення. Пізніше Клайд зрозумів, що це їм мало допомагало.

Побудова речень навколо якогось елемента, який є свого роду його смисловим центром, характерне для стилю коротких повідомлень і кореспонденцій англійської преси. Автори таких повідомлень намагаються включити в одне речення якомога більше деталей:

Thousands of Algerians tonight fled from the "dead city" of Orleansville after a twelvesecond earthquake had ripped through Central Algeria, killing an estimated one thousand one hundred people. Смисловим центром цього речення є підмет підрядного речення (*earthquake*), а не підмет головного речення (*thousands of Algerians*). Навколо цього смислового центру групуються всі інші елементи речення – відбувся землетрус, як довго він тривав, скільки людей загинули. При перекладі це речення краще розбити на два або навіть три незалежних речень: Сьогодні вночі тисячі алжирців втекли від землетрусу із "мертвого міста" Орлеансвіля. Землетрус, який тривав дванадцять секунд, охопив центральні райони Алжиру. Як вважають, загинули тисяча сто чоловік.

Прийом об'єднання двох і більше речень оригіналу в одне речення при перекладі застосовується як для уточнення змісту, так і з стилістичною метою. Найчастіше до нього доводиться вдаватися, коли зв'язок між сусідніми реченнями настільки тісний, що їх формальне об'єднання диктується побудовою української фрази. Об'єднання речень при перекладі може бути доцільним, коли жодне з них не містить закінченої думки:

The Chartists had not planned to assemble in arms on Kennington Common. Or march thence to the Houses of Parliament. Чартисти ї не думали про те, щоб прийти у Кеннінгтонський парк із зброєю в руках і організувати похід до парламенту. У цьому випадку немає підстав залишати логічно тісно зв'язані між собою думки розділеними. Розчленування і об'єднання речень при перекладі тісно пов'язані із вибором синтаксичних конструкцій.

Лекція 8.

Розпізнавання усних висловлювань. Особливості фонематичного декодування сигналів

- 1) Членування мовного потоку.
- 2) Поняття «фраза», «такт», «склад».
- 3) Система "текст-в-мовлення".

Розрізняють *розпізнавання машинної мови* і *перетворення її на текст*, її використання у вигляді команд і справжнє розуміння змісту, як це робить людина. Останнє вимагає знання усього обсягу стилістичних та семантичних конструкцій, правил використання слів та висловів.

Людська мова – це серія звуків. А звук – це поєднання хвиль, у яких дві характеристики: амплітуда і частота.

Мовленнєвий потік членується на відрізки різного розміру. Ці відрізки і є основними фонетичними одиницями. До них належать фрази, такти, склади і звуки. Найкрупнішою фонетичною одиницею є фраза. **Фраза** – це відрізок мовлення, що становить собою інтонаційно-змістову єдність, виділену з двох боків паузами. У тому самому значенні, що й фраза, вживається термін синтагма. Кожна фраза (фонетична синтагма) пов'язана зі змістом та інтонаційним малюнком і має власний (фразовий) наголос. У середині фрази не може бути паузи, оскільки вона б спотворила зміст. Фраза розпадається на такти.

Такт – частина фрази, об'єднана одним наголосом. У мовознавстві такт ще називають фонетичним словом, оскільки службові слова, як правило, не мають наголосу і примикають до повнозначного слова з наголосом; нерідко і повнозначне слово втрачає наголос, передаючи його

попередньому чи наступному слову. Такти членуються на склади, оскільки кожен такт складається з декількох зростань і спадів звучності.

Склад у свою чергу складається зі звуків – мінімальних фонетичних одиниць. Потрібно зазначити, що в мовленні звуки, як правило, окремо не вимовляються, а лише у складі, який є найменшою одиницею мовленнєвого ланцюжка.

У процесі перекладу усного мовлення виникає низка складних ситуацій, адже варто звернути увагу й на те, що люди висловлюють свою думку, як їм заманеться, не звертаючи уваги на те, наскільки це грамотно і літературно. Іноді навіть те, що висловилося вголос, не відповідає тому, що малося на увазі, тому навчити машину розуміти людей – задача неймовірно складна.

Звук представляється в пам'яті комп'ютера як масив значень, що показують коливання амплітуди за часом.

У літературі з мовленнєвих технологій терміни "переведення мовлення в текст", "переведення тексту в мову" і "переведення мовлення в мову" відносять до систем, що включають перетворення тексту й мови, коли на кожному введенні й на кожному виведенні є різні мови. Переклад мовлення в текст означає переклад мовлення однією мовою (джерело введення) в текст іншою мовою (результат виводу), коли сказане слово перетворюється на текст для осіб, у яких є проблеми зі слухом або рухливістю. Прикладами систем такого перекладу є IBM ViaVoice і Dragon NaturallySpeaking - обидві перекладають між сімома мовами.

Приклад застосування технології перетворення мовлення на текст - це перетворення усної мови на телебаченні на "закадрові титри" для глухих і тих, хто погано чує. це перетворення – не переклад, але все ж таки має чітку зв'язаність і з людською діяльністю, і з інтерпретацією. закадрові субтитри є письмовим варіантом не лише того, що мовиться, а й коментарем звуків, наприклад, "телефонує телефон" або "звуки кроків" для зручності людей,

які мають проблеми зі слухом. Глядач може вмикати субтитри за допомогою спеціального декодера.

Система "текст-в-мовлення" вмикається, коли текст (джерело введення) перекладається на мовлення (результат виведення), тобто текст перетворюється на голосове повідомлення переважно для тих, хто має проблеми із зором.

Прикладами систем перекладу тексту в мовлення є «Talking Translator Pro» фірми «AbhiSoft Technologies», що перекладає між вісьмома мовами, та «A T&T Labs Natural Voices», яка перекладає між п'ятьма мовами.

За технологією "мовлення-в-мовлення" мовлення (джерело вводу) перекладається в мовлення (результат виводу) для дво- та одномовних користувачів. Серед систем перекладу мовлення в мовлення широко відомі «M ASTO R» (Multilingual Autom atic Speech to Speech Translator) фірми IB M, призначена для встановлення мовленнєвого спілкування між особами, що не володіють спільною мовою. Функція дво- і багатомовної систем перекладу мовлення в мовлення аналогічна функції усного перекладача, наприклад, на ділових переговорах.

Більшість сучасних систем розпізнавання побудовані на базі рекурентних нейромереж (RNN). RNN можуть розуміти контекст речення завдяки внутрішній пам'яті та зворотним зв'язкам. Це дає змогу передбачати слова в реченнях, враховуючи попередні слова. Виглядає це так: звукова хвиля дробиться на фрейми, після чого з них витягують найважливіші характеристики. Потім нейронна мережа дає прогноз і вгадує ймовірність того, яка фонема прозвучала на фреймі. Декодування видає гіпотези. Наприклад, вираз "Тарас Шевченко – український поет" набагато ймовірніший, ніж "Тарас Шевченко – український поем".

Рекурентна нейронна мережа вчиться передбачати фонему для висловлювання на основі ідентифікації різних фонем. Нині інтенсивно проводяться дослідження в галузі перекладу усної мови. Компанія «IBM»,

яка вважається лідером у цій галузі, випустила програмний пакет «ViaVoice» 4.1, який дає змогу комп'ютеру сприймати до 140 слів на хвилину неперервного диктування. Попередні версії цієї програми

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Кальниченко, О.А. & Черноватий, Л.М. (2020). *Енциклопедія перекладознавства*. Вінниця: Нова Книга.
2. Колесниченко, А.В. & Жмаєва, Н.С. (2018). Граматичні труднощі автоматизованого перекладу науково-технічної літератури. *Науковий вісник ПНПУ ім. К. Д. Ушинського*, 27, 134–141.
3. Сидоренко, Ю.І. (2017). *Практикум з машинного перекладу: навчальний посібник*. Миколаїв.
4. Alonso, E. & Calvo, E. (2015). Developing a Blueprint for a Technology-mediated Approach to Translation Studies. *Meta*, 60(1), 135–157. <https://doi.org/10.7202/1032403a>
5. Bashmanivskiy, O. (2018). Проблеми автоматизованого перекладу ділової кореспонденції за допомогою вільних програмних продуктів. *Society. Document. Communication. Соціум. Документ. Комунікація*, (2), 79–90. <https://sdc-journal.com/index.php/journal/article/view/98>
6. Burchardt, A., Lommel, A., Bywood, L., Harris, K. & Popović, M. (2016). Machine translation quality in an audiovisual context. *Target*, 28(2), 206–221. <https://doi.org/10.1075/target.28.2.03bur>
7. Cid, C. G., Colominas, C., Oliver, A. (2020). Language industry views on the profile of the posteditor. *Translation Spaces*, 9(2), 283–313. <https://doi.org/10.1075/ts.19010.cid>
8. Cifuentes-Ferez & Rojo, P. A. (2015). Thinking for Translating. Think-aloud Protocol on the translation of manner-of-motion verbs. *Target*, 27(2), 273–300. <https://doi.org/10.1075/target.27.2.05cif>

9. EMT Expert Group (2009). *Competences for professional translators, experts in multilingual and multimedia communication*. Brussels.
10. Esqueda, M. D. (2021). Machine translation: teaching and learning issues. *Trabalhos em Linguística Aplicada*, 60(1). 282–299.
11. Jisun, Sh., Kim E. (2017). The Emergence of an Artificial Intelligence Translation System. *The Journal of Translation Studies*, 18(5), 91–110.
12. Killman, J. (2018). Translating the same text twice. *The Journal of Internationalization and Localization*, 5(2), 114–141.
<https://doi.org/10.1075/jial.18003.kil>
13. Vieira L. N. & Alonso E. (2020) Translating perceptions and managing expectations: an analysis of management and production perspectives on machine translation. *Perspectives*, 28(2), 163–184.
<https://doi.org/10.1080/0907676X.2019.1646776>