

М. В. Давидов, О. В. Лозинська, В. В. Пасічник
Національний університет "Львівська політехніка",
кафедра інформаційних систем та мереж

ІНФОЛОГІЧНА МОДЕЛЬ КОНЦЕПТІВ УКРАЇНСЬКОЇ ЖЕСТОВОЇ МОВИ

© Давидов М. В., Лозинська О. В., Пасічник В. В., 2014

Розроблено інфологічну модель концептів для побудови системи перекладу української жестової мови. Досліджено адекватність інфологічної моделі на корпусі паралельних речень «українська словесна мова – українська жести мови». Оцінено якість системи перекладу з використанням інфологічної моделі концептів зменшення відсотка помилково перекладених слів у тих реченнях, які містили концепти, з 42,83 % до 34,58 % WER.

Ключові слова: українська жести мова, корпус паралельних текстів, концепт, інфологічна модель, граматики.

Entity-relationship model of concepts for building a translation system of the Ukrainian Sign Language is developed. Adequacy of the entity-relationship model of parallel sentences "Ukrainian Spoken Language - Ukrainian Sign Language" is investigated. The evaluation quality of the translation system using the entity-relationship model concepts shows a reduction in the percentage of incorrectly translated words in those sentences that contained concepts, from 42,83 % to 34,58 % WER and from 18,35 % to 14,65 % PER.

Key words: Ukrainian Sign Language, bilingual corpora, concept, entity-relationship model, grammar.

Вступ

Засоби автоматичного перекладу допомагають людям спілкуватися й долати мовні та культурні бар'єри. Одним із найвідоміших засобів перекладу сьогодні є Google Translate, який поєднує системи правил та статистичні методи для перекладу низкою мов світу. Проте ще не вирішено проблему перекладу жестовою мовою (ЖМ) і навпаки, що обмежує у спілкуванні людей, які володіють лише жестовою мовою. Жести мова – це візуально-просторова мова, яка має власну організацію та використовується спільноту людей із вадами слуху. У жестовій мові для передавання інформації застосовують жести рук, артикуляцію губ, міміку. Українська жести мова (УЖМ), як й інші відомі жести мови, має свою граматику і правила, відмінні від української словесної мови (УСМ).

У світі не існує єдиної універсальної жестової мови. Жести мови різних країн розвивалися незалежно, а тому значно відрізняються одна від однієї. Крім того, кожна країна має місцеву жести мову або різноманіття мов, що відображають культуру людей із вадами слуху, що там живуть. Наприклад, у Швейцарії існує чотири місцеві жести мови, якими користуються глухі французької, італійської, німецької та швейцарської громад [1]. Попри використання єдиної словесної мови (СМ) в Німеччині та Австрії, австрійська жести мова (Austrian Sign Language, ÖGS) та німецька жести мова (German Sign Language, DGS) є двома не пов'язаними між собою мовами (крім того, у Німеччині нараховується кілька діалектів DGS). Подібна ситуація у США та Великобританії: словесна офіційна мова цих країн англійська. Однак люди із вадами слуху в Америці спілкуються американською жестовою мовою – амсленом (ASL), а у Великобританії – британською жестовою мовою (BSL) [2].

Складність розроблення системи автоматичного перекладу української жестової мови посилюється відсутністю великих словників та корпусів УЖМ [3]. Оскільки жести мови використовують для подання інформації різні жести і вирази обличчя замість звуків, то для

адаптування систем перекладу писемних мов для перекладу жестової мови необхідно розробити систему запису всіх елементів жестової мови, створити відповідні перекладацькі словники й системи граматичних правил для розуміння жестової мови. Крім того, значна лінгвістична відмінність жестової мови від словесної ускладнює процес перекладу.

У відомих системах машинного перекладу жестових мов, оснований на статистичних моделях, встановлюються прямі відношення між жестами та словами мови. Такий підхід може забезпечити якісний переклад лише за умови використання значних навчальних корпусів жестової і словесної мов. Відсутність таких корпусів для української жестової мови вимагає застосування альтернативного підходу – введення концептів та дослідження відношення між цими концептами.

Постановка проблеми

Під концептом розуміємо значення жестового або словесного висловлювань, які виражають те саме поняття (процес, дію, ознаку) і можуть бути скорочені лише у випадку, коли слово або жест відновлюється з контексту. Наприклад, словосполучення «годинник йде» вважаємо концептом, тому що в жестовій мові для показу цього словосполучення використовують один жест «ГОДИННИК_ЙДЕ», а словосполучення «людина йде» не вважатимемо концептом (використовують жести «ЛЮДИНА» та «ЙТИ»).

Людина, яка вільно володіє обома мовами, швидко розуміє зміст речень і наявних у реченні концептів та перефразовує із застосуванням засобів мови перекладу. Така людина є експертом, який володіє необхідними для перекладу знаннями. Комп'ютерна програма, яка може правильно автоматично перекладати концепти, повинна мати засоби для адекватного подання знань експерта у вигляді відповідної бази даних, яка містить інформацію про концепти та їх переклад.

У статті описано інфологічну модель концептів, зорієнтовану на переклад української жестової мови. Сутність інфологічного моделювання полягає у виокремленні інформаційних об'єктів предметної галузі, які підлягають зберіганню в базі даних, а також у визначенні характеристик об'єктів і зв'язків між ними. Характеристиками чи властивостями об'єктів є атрибути. Інфологічну модель побудовано без урахування засобів і технологій реалізації проекту [4].

На рис. 1 зображено структуру системи перекладу з української жестової мови на українську словесну та навпаки.

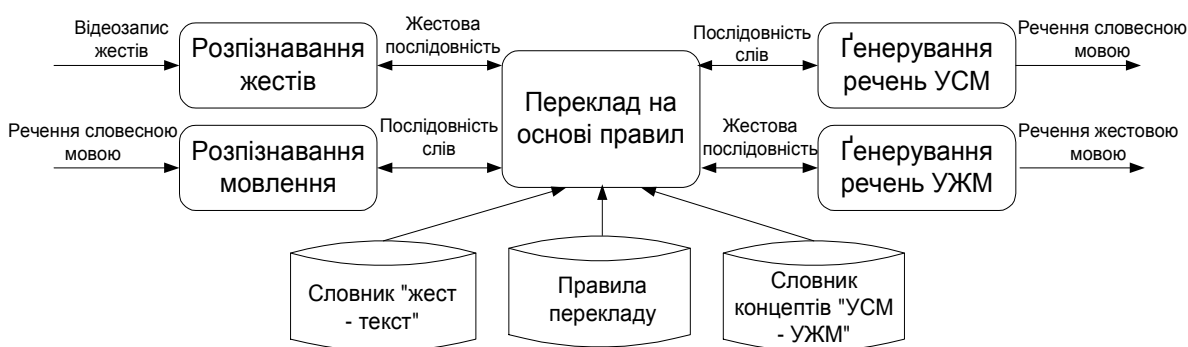


Рис. 1. Компоненти системи перекладу української жестової мови

На вхід процесу перекладу подається послідовність жестів у вигляді записаних глосів [2] (один глос відповідає одному жесту) або послідовність слів залежно від того, з якої мови і якою здійснюється переклад. Для запису жестів у вигляді глосів використано слова, записані великими літерами. Наприклад, слово “бачити” українською жестовою мовою записано таким чином: “БАЧИТИ”. Модуль машинного перекладу на основі правил трансформує вхідну послідовність жестів у вихідну послідовність слів української мови. Цей модуль здійснює процес перекладу за допомогою спеціальних правил перекладу, словника «жест-текст» та словника концептів «УСМ – УЖМ». На виході процесу отримуємо речення українською словесною мовою. Система машинного перекладу також здійснює зворотний переклад, тобто з української словесної українською жестовою мовою.

Інфологічна модель розробляється на основі аналізу лінгвістичних відношень між концептами українською словесною та українською жестовою мовами. Адекватність інфологічної моделі досліджено з використанням корпусу паралельних речень «УСМ-УЖМ». Оцінено якість системи перекладу з використанням інфологічної моделі концептів.

Аналіз останніх досліджень та публікацій

Від 1960 р. досліджують жестові мови не тільки в США [5], але й у Європі [6], Африці [7] та Азії [8]. В Україні досліджують українську жестову мову Р. Г. Краєвський [9], Н. Б. Адамюк, І. І. Чепчина [10].

У системах перекладу жестової мови [2,11,12] для експериментів використовуються корпуси паралельних текстів із певної предметної галузі.

Переклад концептів наявний у тій чи іншій формі у працях [13, 14]. У праці іспанських учених Р. Сан-Сегундо та ін. (R. San-Segundo et al.) для перекладу іспанської жестової мови проаналізовано відношення між концептами іспанської словесної та жестової мов. На основі цих відношень експерт описав основні правила перекладу. Розроблено модуль машинного перекладу на основі правил. Процес перекладу здійснюється в два етапи. На першому етапі кожне слово зіставляється з одним або декількома синтаксичними тегами. Після цього, враховуючи правила перекладу, позначені слова трансформуються в жести відповідно до відношень між концептами. Модуль перекладу на основі правил містить 153 правила. Для оцінювання результату перекладу використано оцінку помилок перекладу жестів SER (Sign Error Rate), яка становить 31,60 %.

С. Балдассаррі та ін. [15] розробили систему перекладу іспанського тексту іспанською жестовою мовою. Модуль машинного перекладу на основі правил використовує синтаксичні та морфологічні характеристики слів та їхнє семантичне значення для генерування відповідних жестів. Систему протестовано на 92 реченнях, які містять 561 слово. Результат перекладу становить 96 % правильно перекладених слів.

У праці українських вчених Ю. В. Крака та ін. [3] розглянуто проблему побудови системи автоматизованого перекладу тексту українською мовою українською жестовою мовою. Учені розробили інформаційну технологію, основу на узагальнених граматичних конструкціях простих речень українською та жестовою мовами. Отримано 764 таких узагальнених граматичних конструкцій із 6559 речень. Створено інфологічну модель бази даних для узагальнених граматичних конструкцій УСМ та УЖМ у вигляді таблиць та зв'язків між ними. У праці [15] запропоновано практичну реалізацію інфологічних моделей для словників флективної та жестової мов для роботи із цими граматичними конструкціями й алгоритм перекладу. Описано інфологічні подання словників української мови як множини інфінітивів, флексій та параметрів словозміни слів і жестової мови як множини жестів та параметрів, що їх характеризують. За допомогою запропонованої інформаційної технології отримано однозначний переклад без спотворення змісту в 100 % речень, узятих із шкільної програми вивчення української мови жестів.

Проте у працях [3, 15] не враховано те, що кожне висловлювання має певне семантичне навантаження. Наприклад, у граматичних конструкціях «я йду» та «час йде» жест «ЙТИ» показують по-різному. Отже, ми не можемо цей жест перекласти однозначно без врахування контексту. Для вирішення цього завдання треба створити словник концептів та їх відповідників в УСМ та УЖМ.

Дослідження відношень між концептами у корпусі «УСМ–УЖМ»

Основним завданням під час перекладу української жестової мови, як і у всіх інших випадках перекладу тексту однією мовою іншою, є правильне передавання змісту тексту, що перекладається. Для цього необхідно повністю розуміти вихідний текст, що є доволі складним завданням [16]. Для вирішення проблем, які виникають під час побудови систем машинного перекладу тексту жестовою мовою та навпаки, важливо визначити відношення між концептами словесної та жестової мов, які передають зміст повідомлення [13]. При цьому треба врахувати, що один жест жестової мови може виражати ціле словосполучення, та навпаки, деякі слова вимагають пояснення у вигляді декількох жестів. Крім того, більшість висловлювань жестовою мовою набувають кількох значень, які треба уточнювати у словесній мові.

Ураховуючи особливості відношень між словами, жестами й концептами, виокремлено чотири основні ситуації, які треба враховувати під час перекладу тексту українською жестовою мовою українською словесною мовою та навпаки.

1. Концепт подається одним словом УСМ та одним жестом УЖМ (табл. 1).

У цьому випадку слово безпосередньо перекладають одним жестом. Переклад простий, тому що одному жесту з УЖМ відповідає одне значення.

Таблиця 1

Одне слово відповідає одному жесту

№	Українська жестова мова	Українська словесна мова
1	Я ЛЮБИТИ ВІНОГРАД ЗЕЛЕНИЙ	Я люблю зелений виноград
2	Я ЖИТИ МІСТО	Я живу в місті

2. Концепт, який подається декількома словами, відповідає одному жесту (табл. 2). Наприклад, висловлювання УСМ «серце б'ється рівно» перекладають за допомогою одного жесту. Приклад зображено на рис. 2.

Таблиця 2

Декілька слів відповідають одному жесту

№	Українська жестова мова	Українська словесна мова
1	Я НЕСТИ_РЮКЗАК	Я несу рюкзак
2	СЕРЦЕБИТТЯ	Серце б'ється рівно



Рис. 2. Зображення жесту «серце б'ється рівно»

3. Концепт, який подають одним словом, показують кількома жестами (табл. 3). Така ситуація виникає в таких випадках:

- 1) переклад дієслів, коли концепт, який відповідає одному дієслову, генерує декілька жестів;
- 2) переклад загальних і специфічних іменників;
- 3) переклад іменника у множині;
- 4) переклад іменника з урахуванням статі.

Під час перекладу дієслів українською жестовою мовою треба враховувати, що в українській словесній мові існують три форми часу: минулий, теперішній та майбутній. Для того, щоб виразити дієслово теперішнього часу, в українській жестовій мові використовують інфінітив, а для вираження дієслів минулого або майбутнього часів використовуються інфінітив та допоміжні жести «був» та «буду» відповідно (табл. 3). Наприклад, концепт УСМ «говорив» відповідає жестам «говорити + був», а концепт «говоритиму» – жестам «говорити + буду».

У жестовій мові трапляються ситуації, коли немає відповідного жесту для певного концепту, який може бути виражений одним словом української словесної мови. Наприклад, в українській жестовій мові є відповідні жести для позначення слів «золото», «срібло», «мідь» тощо, але нема жесту для позначення слова «метал». Те саме стосується жесту для позначення слова «меблі»: є жести для слів «стіл», «стілець», «ліжка». Тому для вирішення цієї проблеми в жестовій мові використовують декілька жестів для позначення загального поняття (табл. 3).

Для перекладу жесту у множині є кілька способів: повторення жесту, використання додаткових жестів «багато», «різноманітні», «декілька» або показ жесту обома руками. Наприклад, слово «цукерки» можна показати такими способами: дворучним жестом «ЦУКЕРКИ», повторенням жесту «ЦУКЕРКА ЦУКЕРКА» чи «ЦУКЕРКА+БАГАТО». Кілька прикладів наведено на рис. 3 [17].



а



б

Рис. 3. Зображення жестів: а – «багато»; б – дворучного жесту «цукерки»

Для позначення статі об'єкта чи суб'єкта до основного жесту додається додатковий жест «чоловік» або «жінка». У більшості випадків подібні уточнення статі робити немає необхідності, тому що це ускладнюватиме переклад. Проте у деяких випадках уточнення статі є необхідним для правильного передавання змісту (табл. 3).

Таблиця 3

Одне слово показують кількома жестами

№	Українська жестова мова	Українська словесна мова
1	Я КУПИТИ БУВ ХЛІБ	Я купила хліб
2	Я (НЕ МАЮ) СТІЛ СТІЛЕЦЬ ЛІЖКО БУДИНОК МІЙ	Я не маю меблів у своєму будинку
3	Я ЛЮБИТИ ПТАШКА ПТАШКА	Я люблю пташок
4	ВІН МАТИ (ЦУКЕРКА 2 РУКАМИ)	Він має цукерки
5	Я БАЧИТИ БАГАТО ДЕРЕВА ЗЕЛЕНИЙ	Я бачу зелені дерева
6	МОЯ БАБУСЯ МАТИ (КОЗА + ЖІНКА)	Моя бабуся має козу

4. Декілька концептів (декілька слів) відповідають декільком жестам (табл. 4).

Згаданий випадок є найскладнішим із погляду перекладу, оскільки треба згенерувати жести УЖМ, які б відповідали за змістом декільком словам словесної мови. Наприклад, дієслово «йти» залежно від суб'єкта дії показують різними жестами: «Дорога йде рівно» і «Час йде швидко». У цьому випадку жест-дієслово залежить від суб'єкта дії.

Таблиця 4

Декілька слів відповідають декільком жестам

№	Українська жестова мова	Українська словесна мова
1	ДОРОГА РІВНО	Дорога йде рівно
2	ЧАС ЛЕТИТЬ	Час йде швидко

Інфологічна модель словника відношень між концептами та їхнім поданням в українській та жестовій мові

На основі аналізу можливих випадків відповідності жестових сполучень, словосполучень і концептів розроблено інфологічну модель словника відношень між концептами. Модель містить такі сутності: «Концепт», «Словосполучення», «Слово у словосполученні», «Граматична категорія слова», «Граматичне значення», «Граматична форма слова», «Початкова форма слова», «Жестове сполучення», «Жест у сполученні», «Граматична категорія жесту», «Граматична форма жесту», «Початкова форма жесту» та подана діаграмою «сутність–зв'язок» (рис. 4) [4].

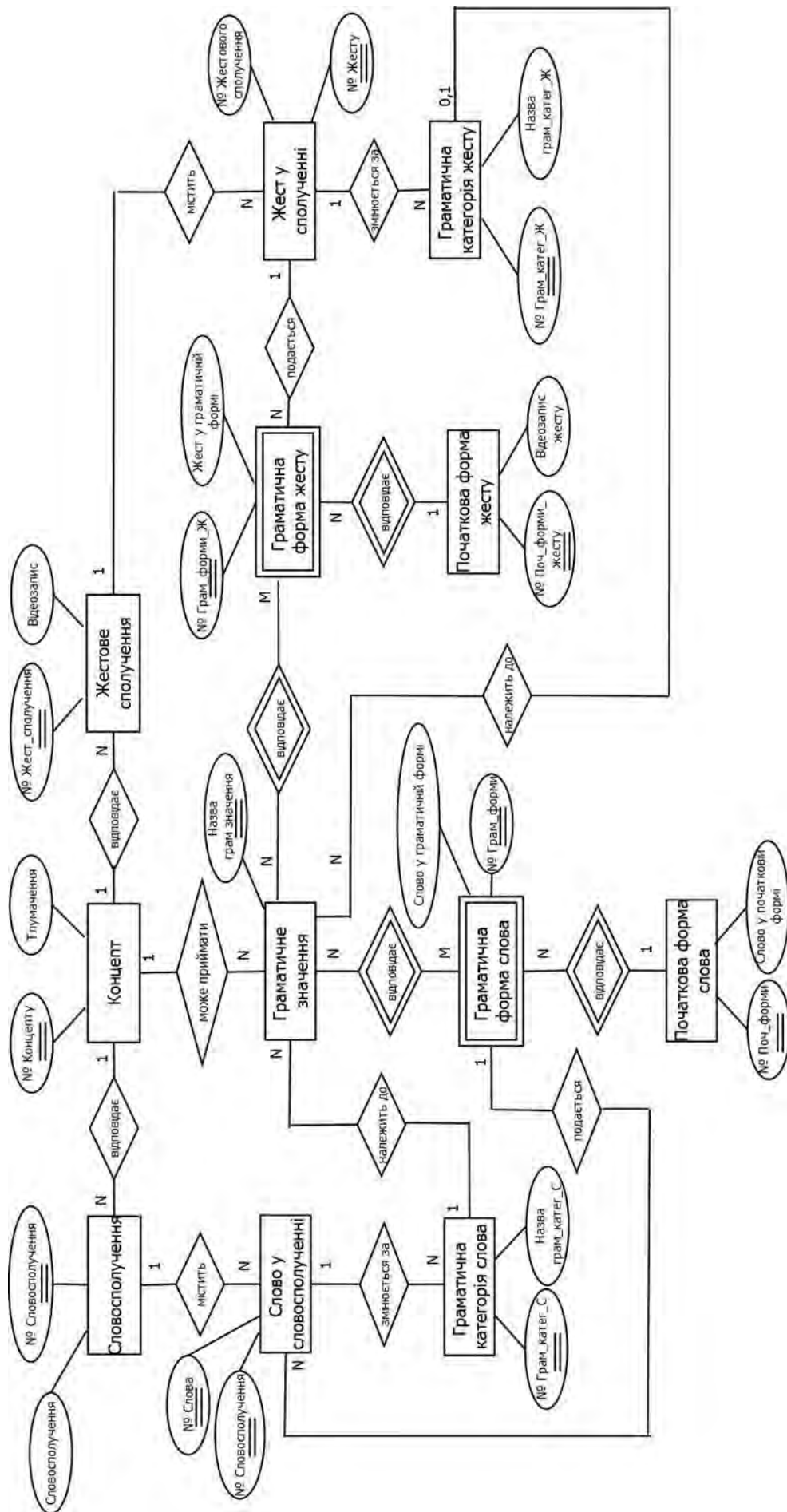


Рис. 4. Інфологічна модель "сутність-зв'язок" словника концептів словесної та жестової мов

Сутність «Концепт» у цій моделі відповідає певному висловлюванню (предмету, дії, ознаці тощо), яке має подання в українській словесній мові та в українській жестовій мові. Ця сутність має такі атрибути: «№ Концепту» (первинний ключ) та «Глумачення» концепту.

Концепт в українській словесній мові виражається одним словом або словосполученням. Одне слово задається як вироджене словосполучення, яке складається з одного слова. Словосполучення – це смислове і граматичне поєднання одного або декількох повнозначних слів. Сутність «Словосполучення» має такі атрибути: «№ Словосполучення» (первинний ключ) та «Словосполучення». Наприклад, перелік значень для відношення Словосполучення (№ Словосполучення, Словосполучення) = {(1, Прийшла осінь), (2, Дівчина з'їла яблуко)}.

Сутність «Слово у словосполученні» містить атрибути «№ Слова» (порядковий номер слова у словосполученні), «№ Словосполучення», які є складеним первинним ключем.

Кожне слово у словосполученні вказано в певній граматичній формі. Граматична форма слова – це написання слова у конкретній формі плюс сукупність граматичних значень, які цьому написанню відповідають [18]. Наприклад, граматична форма слова «зеленого» відповідає граматичним значенням: родовий відмінок, чоловічий рід, число однина. Сутність «Граматична форма слова» містить ключовий атрибут «№ Грам_форми» та атрибут «Слово у граматичній формі».

Різні граматичні форми одного слова відповідають початковій формі слова. Початкова форма відповідає таким граматичним значенням: називному відмінку однини – для іменників та інфінітиву – для дієслів. Сутність «Початкова форма слова» містить ключовий атрибут «№ Поч_форми» та атрибут «Слово в початковій формі».

Граматична форма слова може характеризуватися одним чи кількома граматичними значеннями. Граматичне значення – узагальнене (абстрактне) мовне значення, яке властиве групам слів, словоформ, синтаксичних конструкцій. Наприклад, різні за своїм значенням іменники «розум», «дуб», «хлопець» мають однакові граматичні значення: називного відмінка, числа однини, виражають належність до чоловічого роду. Сутність «Граматичне значення» містить атрибут «Назва грам_значення», що є первинним ключем.

Споріднені граматичні значення об'єднуються у граматичні категорії. Наприклад, чоловічий, жіночий і середній рід об'єднуються в одну категорію роду.

Деякі граматичні категорії мають і граматичне і семантичне значення, а деякі мають лише граматичне. При цьому така семантична властивість може залежати й від самого слова. Наприклад, рід для живих істот ссавців має семантичне значення, а для неживих предметів – лише граматичне.

Граматичні категорії можуть бути спільними для декількох частин мови. Категорія числа охоплює багато частин мови (іменник, прикметник, займенник, дієслово), так само й категорія роду, а категорії особи, часу, виду, способу властиві тільки дієсловам, що пояснюється самим виявом предметів і явищ. Граматична категорія своїх показників не має, вона виявляється через граматичні значення, засобами вираження яких є граматичні форми. Сутність «Граматична категорія слова» містить ключовий атрибут «№ Грам_катег_С» та атрибут «Назва грам_катег_С». Наприклад, перелік значень для відношення Граматична категорія слова (№ Грам_катег_С, Назва грам_катег_С) = {(1, Відмінок), (2, Число), (3, Рід), (4, Час), (5, Вид), (6, спосіб), (7, Особа)}.

У жестовій мові концепт виражається одним або декількома синонімічними «Жестовими сполученнями». Жестове сполучення має два атрибути: «№ Жест_сполучення» (первинний ключ) та «Відеозапис» жестового сполучення. Жестове сполучення може складатися з одного або декількох жестів. Наприклад, перелік значень для відношення Жестове сполучення (№ Жест_сполучення, Відеозапис) = {(1, Відеозапис_1), (2, Відеозапис_2)}. Сутність «Жест у сполученні» містить ключовий атрибут «№ Жесту» та атрибут «Жест» (запис жесту у вигляді глоса).

Граматична категорія жесту. В українській мові такі граматичні категорії, як рід, число, відмінок, час і особа передаються за допомогою закінчень. У жестовій мові немає закінчень, які у словесній мові передають відтінки думок, час, особу тощо, проте граматичні категорії є. Це категорії числа, роду – для іменників та категорії часу, виду, способу – для дієслів. Сутність «Граматична категорія жесту» містить ключовий атрибут «№ Грам_катег_Ж» та атрибут «Назва грам_катег_Ж». Сутність «Початкова форма жесту» має такі атрибути «№ Поч_форми_жесту» (первинний ключ) та «Відеозапис» показу жесту.

Отже, під час перекладу українською жестовою мовою враховується залежність граматичного значення жесту від граматичного значення слова.

Аналіз результатів

Для тестування системи машинного перекладу використано граматичний парсер [19], який буде дерево граматичного розбору речення. Для перекладу української жестової мови застосовуються граматичні правила перекладу, описані у роботі [20], та проводиться пошук для заміни кожного слова (словосполучення) українською словесною мовою відповідним жестом із використанням інфологічної моделі концептів.

Приклад дерева синтаксичного розбору речення «Хлопець несе рюкзак зі школи» наведено на рис. 5.

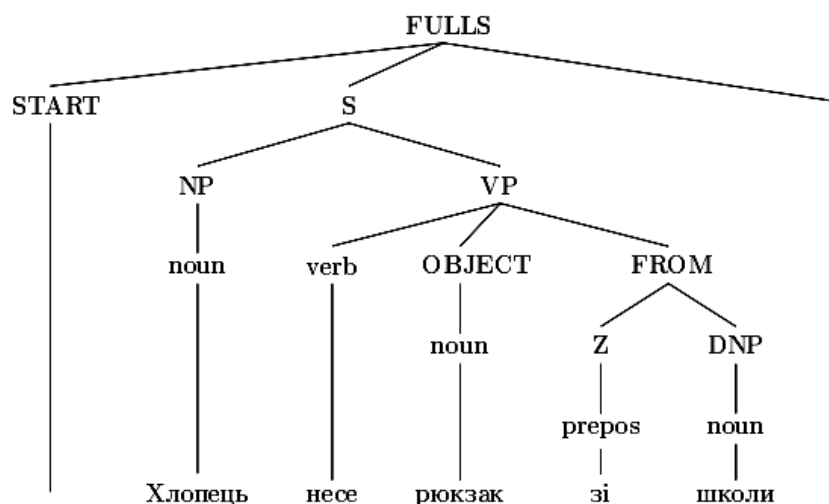


Рис 5. Дерево граматичного розбору речення «Хлопець несе рюкзак зі школи»
 Позначення: FULLS – речення; S – основа речення; VP – група дієслова;
 NP – підмет; DNP – додаток

Розгляньмо, як здійснюється переклад з використанням концептів для наведеного речення (табл. 5). Слово «хлопець» відповідає жесту «ХЛОПЕЦЬ», а словосполучення з двох слів «несе рюкзак» є концептом, який перекладається одним жестом «НЕСТИ_РЮКЗАК». Оскільки в жестовій мові прийменники не позначають (їх вживають лише в калькованій жестовій мові), то ми його не перекладаємо. Щоб розрізнити словосполучення «зі школи» і «до школи», в жестовій мові вказує напрямку руху. Слово «школа» має відповідник у жестовій мові «ШКОЛА». Зрозуміло, що якщо кожне словосполучення «зі школи», «з магазину», «з бібліотеки», тощо, «до школи», «з магазину», «з бібліотеки» тощо записувати окремо, то ми отримаємо великий набір однакових за структурою словосполучень, які можна виокремити в групу «зі (до) школи (магазину, бібліотеки і т.ін.)». Виокремлення таких груп можна використати для розроблення онтології з граматичною компонентою для українських жестової та словесної мов.

Таблиця 5

Переклад із використанням концептів

№	Українська жестова мова	Українська словесна мова
1	ХЛОПЕЦЬ	Хлопець
6	НЕСТИ_РЮКЗАК	Несе рюкзак
10	ШКОЛА	Зі школи

Для тестування системи машинного перекладу УЖМ на основі правил із використанням словника концептів взято речення з корпусу паралельних текстів «українська словесна мова – Українська жестова мова», який містить 220 простих речень різної тематики українською словесною мовою та їх перекладів жестовою, 932 слова УСМ і 839 слова УЖМ. Основні характеристики корпусу наведено в табл. 6. Словник концептів розроблено згідно з описаними відношеннями між концептами та містить 50 жестових концептів та їх відповідних перекладів українською словесною мовою. Фрагмент словника концептів «УСМ–УЖМ» наведено в табл. 7.

Статистика корпусу паралельних текстів «УСМ–УЖМ»

	УСМ	УЖМ
Кількість речень	220	
Кількість слів	932	839
Словник	432	398

Таблиця 7

Фрагмент словника концептів «УСМ–УЖМ»

№	Українська жестова мова	Українська словесна мова
1	КОЗА + ЖІНКА	Коза
2	СТІЛ + СТІЛЕЦЬ + ЛІЖКО	Меблі
3	БАГАТО + ЯБЛУКО	Яблука
4	ДОРОГА РІВНО	Дорога йде рівно
5	СЕРЦЕБИТТЯ	Серце б'ється швидко
6	НЕСТИ_РЮКЗАК	Несу рюкзак
7	ЧАС_ЛЕТИТЬ	Час летить
8	ЛІГАТИ	Лігати

Для оцінювання якості системи перекладу з української словесної українською жестовою мовою використано оцінки (word error rate) – оцінка помилок перекладу, PER (position-independent word error rate) – позиційно-незалежна оцінка помилок. Результати оцінювання якості перекладу без використання словника концептів та з його використанням наведено у табл. 8. Використання двомовного словника концептів «УСМ–УЖМ» та правил визначення, до якого концепту належить слово в певному контексті української жестової мови, дало змогу покращити результати системи машинного перекладу на основі правил з 42,83 % до 34,58 % WER (що менша WER, то кращий результат). Переклад покращився на 40 реченнях, які містили концепти. Наприклад, речення «Серце б'ється швидко» з використанням словника концептів перекладено на УЖМ «СЕРЦЕБИТТЯ» (швидкий темп жесту), а без використання словника концептів – «СЕРЦЕ БИТИ ШВИДКО» (у цьому випадку жест «БИТИ» означав «завдавати ударів»), тобто повністю втрачено зміст речення.

Таблиця 8

Оцінка якості перекладу

	WER, %	PER, %
УСМ→УЖМ (без використання словника концептів)	42,83	18,35
УСМ→УЖМ (з використання словника концептів)	34,58	14,65

Проведені експерименти доводять, що для якісного перекладу української жестової мови належить укласти перелік концептів (деколи концепти становлять ціле речення). Для зменшення відсотка помилково перекладених слів у реченні необхідно згрупувати концепти та розробити на їх основі онтологію українських словесної та жестової мов.

Висновки

На основі аналізу лінгвістичних відношень між концептами українською мовою й українською жестовою мовою розроблено інфологічну модель концептів для побудови системи перекладу української жестової мови. Протестовано роботу цієї моделі на прикладі корпусу паралельних речень «УСМ–УЖМ». Оцінено якість системи перекладу з використанням інфологічної моделі концептів показало покращення результату перекладу на тих реченнях, які містили концепти, з 42,83 % до 34,58 % WER. Недоліком системи є невелика кількість даних для тестування (220 речень).

У результаті дослідження для перекладу української жестової мови згенеровано граматично правильні конструкції словесної мови, що відтворюють зміст жестового висловлювання. На основі проведених експериментів досліджено, що для якісного перекладу української жестової мови

необхідний перелік концептів, тому доцільно ці концепти згрупувати та розробити на їхній основі онтологію українських словесної та жестової мов. Розроблення такої онтології з граматичною компонентою та використання її в системі машинного перекладу української жестової мови стане темою подальших досліджень.

1. Stone D. *Interpreting in international sign: Decisions of deaf and non-deaf interpreters* / D. Stone, D. Russel // *Paper presented at the World Association of the Sign Language Interpreters in Durban*. – South Africa, 2011.
2. Stein D. *Sign language machine translation overkill* / D. Stein, C. Schmidt, H. Ney // *International Workshop on Spoken Language Translation (IWSLT)*. – Paris, 2010. – P. 337–344.
3. Бармак О. *Комп'ютерна технологія для перекладу з української на жестову мову* / Бармак О., Романишин С. // *Вісник ТНТУ*. – 2012. – Т. 65. – № 1. – С.195–208.
4. Берко А. Ю. *Системи баз даних та знань. Книга 1. Організація баз даних та знань: навч. посібник* / А. Ю. Берко, О. М. Верес, В. В. Пасічник. – Львів: «Магнолія 2006», 2008. – 456 с.
5. Stokoe W. C. *Sign language structure: An outline of the Visual Communication Systems of the American deaf studies in linguistics: Occasional papers* / W. C. Stokoe. – Buffalo: Dept. of Anthropology and Linguistics, University of Buffalo, 1960. – 78 p.
6. Engberg-Pedersen E. *Space in Danish sign language: The semantics and morphosyntax of the use of space in a visual language* / E. Engberg-Pedersen. – Hamburg: Signum Press, 1993. – P. 231–244.
7. Nyst V. *The phonology of name signs: a comparison between the sign languages of Uganda, Mali, Adamorobe and The Netherlands* / V. Nyst, A. Baker // *Cross-linguistic perspectives in sign language research*. – Hambur, 2003. – P. 71–80.
8. Abdel-Fattah Mahmood A. *Arabic sign language: A perspective* / A. Abdel-Fattah Mahmood // *Journal of Deaf Studies and Deaf Education*. – 2005. – 10. – P. 212–221.
9. Краєвський П. Г. *Мова жестів глухих* / П. Г. Краєвський. – К.: Рад. Школа, 1964. – 219 с.
10. Адамюк Н. Б. *Синтаксичні особливості УЖМ: на прикладі простого речення* / Н. Б. Адамюк, І. І. Чепчина // *Жестова мова й сучасність: зб. наук. праць*. – Вип. 4. – К.: Педагогічна думка, 2009. – С. 170–191.
11. Dreuw P. *Speech recognition techniques for a sign language recognition system* / P. Dreuw, D. Rybach, T. Deselaers, M. Zahedi, H. Ney // *Eighth Annual Conference of the International Speech Communication Association*. – 2007.
12. Bungeroth J. *Statistical sign language translation* / J. Bungeroth, H. Ney // *Proc. LREC 2004, Workshop proceedings: Representation and Processing of Sign Languages*. – Lisbon, Portugal, 2004. – P. 105–108.
13. San-Segundo R. *Speech to sign language translation system for Spanish* / R. San-Segundo, R. Barra, R. Córdoba, et al. // *Speech Communication (ISSN: 0167-6393)*. – Vol. 50. – 2008. – P. 1009–1020.
14. Baldassarri S. *An automatic rule-based translation system to Spanish Sign Language* / Baldassarri S., Royo-Santas F. // *New Trends on Human-Computer Interaction*. – 2009. – P. 1–11.
15. Крак Ю. В. *Реализация информационной технологии автоматизированного перевода с флективных языков на жестовый* / Ю. В. Крак, А. В. Бармак, С. А. Романишин // *Материалы IV Международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем»*. – Минск, 2014. – С. 21–30.
16. Лурья А. Р. *Понимание компонентов речевого высказывания* // *Язык и сознание* / Ред. Е. Д. Хомской. – М.: Изд. МГУ, 1979. – С. 217–226.
17. Проект «Spreadthesign» [Електронний ресурс]. – Режим доступу: <http://www.spreadthesign.com/ua/>.
18. Вихованець І. Р. *Грамматика української мови. Синтаксис* / І. Р. Вихованець. – К., 1993. – 368 с.
19. Проект «UkrParser» [Електронний ресурс]. – Режим доступу: <https://github.com/mdavydov/UkrParser>.
20. Лозинська О. В. *Машинний переклад на основі правил для перекладу на українську жестову мову* / О. В. Лозинська, М. В. Давидов, В. В. Пасічник // *Інформаційні технології та комп'ютерна інженерія: Міжнародний науково-технічний журнал*. – Вінниця, 2014. – С. 11–17.