

Лихошерстов Д.О.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Лебедев Д.Ю.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЦИФРОВИХ ЗАСОБІВ СУРДОПЕРЕКЛАДУ

У останні роки відбувається суттєвий прогрес в розробці, адаптації та впровадженні систем сурдоперекладу в суспільство для спрощення життя людей із вадами слуху та мовлення.

У статті проводиться порівняльний аналіз систем сурдоперекладу, спрямований на визначення переваг та недоліків кожної системи для покращення комунікації для людей з вадами слуху. Виконано аналіз актуальних досліджень та підходів, що були представлені в різноманітних додатках, плагінах чи розширеннях для додатків щодо функціональних можливостей актуальних систем сурдоперекладу на базі смартфонів, персональних комп'ютерів та окремих апаратно-технічних рішень. В ході досліджень було проаналізовано ряд рішень, що використовують концептуально різні підходи в перекладі жестів. Найпопулярнішим залишається підхід у використанні живих сурдоперекладачів, котрі можуть бути як і професійними спеціалістами, так і волонтерами. Актуальним рішенням є застосування готових баз жестів для найпопулярніших жестових мов, де є можливість здійснювати переклад для не досвідчених користувачів. Найсучаснішим підходом є застосування технологій розпізнавання зображення на відео та технологій штучного інтелекту.

На жаль, реальна кількість унікальних рішень і відповідно програмних реалізацій у вигляді додатків не дозволяє в повному обсязі оцінити проблеми сурдоперекладу, але відповідно до поточних досліджень це і не є основним критерієм дослідження.

Ключові слова: сурдопереклад, жести, глохонімія, смартфон, додаток, штучний інтелект, порівняльний аналіз.

Постановка проблеми. Відповідно до статистики Всесвітньої організації охорони здоров'я (ВООЗ), опублікованої в пресреліз за 2021/2022 рік, кожен четвертий мешканець планети до 2050 року буде страждати проблемами порушення слуху, а це майже 2.5 млрд людей по всьому світі [1]. Також, відповідно до результатів досліджень ВООЗ, як мінімум 700 млн. людей [2] потребують кваліфікованого медичного догляду органів слуху, а також різноманітних реабілітаційних послуг.

Постійні погіршення слуху призводять до повної або часткової повної глухоти, що зумовлюють негативні процеси в мовленні та відповідно в здатності людини повноцінно комунікувати в суспільстві та бути частиною суспільства.

Наявність великої кількості глухонімих людей у світі створила актуальну проблему, котру намагаються вирішити великою кількістю наукових досліджень. В результаті досліджень формуються практичні інструменти подолання проблем кому-

нікації глухонімих людей в суспільстві. Основним підходом у розв'язання даної проблеми є створення програмних рішень.

Аналіз останніх досліджень і публікацій.

Слід зазначити, що сурдопереклад складається з двох основних частин: міміко-жестового мовлення та дактилології – пальцевої (ручної) абетки, за допомогою якої під час невербальної комунікації можна передавати новоутворені, специфічні слова, назви, абрєвіатури, прізвища та імена тощо, які не мають відповідного жестового позначення слова. Міміко-жестове мовлення – це система кінетичних знаків, що передають характерологічні особливості предметів, явищ навколишньої дійсності за допомогою визначених структурних елементів. Дактилологія або дактильна абетка – це унормований, стандартний набір відповідно до певної вербальної мови рухових або статичних знаків за участю пальців та кисті руки (рук), який є незмінним для позначення певної літери відповідно даної мови. Нині в суспільстві пара-

лельно використовуються два терміни: «перекладач жестової мови» та «сурдоперекладач». Сурдоперекладач – особа, яка володіє дактильним та міміко-жестовим мовленням і здійснює переклад на дактильне та міміко-жестове мовлення інформаційного повідомлення з будь-якої національної мови або навпаки. Внаслідок того, що в різних країнах застосовуються різні невербальні, жести системи спілкування, сурдопереклад, який здобув освіту в одній країні, далеко не завжди виявляється придатний до перекладу в іншій країні, навіть якщо в ній і поширена та ж державна мова. Таким чином, виникає гостра потреба в універсальних засобах сурдоперекладу. Слід зазначити, що під визначенням «універсальний» потрібно розглядати програмний або апаратно-програмний підхід, котрий можна масштабувати (програмно додавати нові регіони/країни і їх жестові мови), вдосконалювати (додавати нові жести/слова для вже наявних мов).

Будь-який підхід сурдоперекладу передбачає, що кожен учасник діалогу повинен володіти повною мірою міміко-жестовим мовленням чи дактилологією, а це, своєю чергою, дуже сильно звужує можливості комунікації глухонімих в суспільстві. Розв'язання даної проблеми є реалізацією програмних рішень у вигляді повноцінних застосунків, сервісів, розширень для смартфонів, персональних комп'ютерів чи спеціальних апаратно-технічних засобів.

Попри великий вибір готових систем сурдоперекладу, існує в них певний набір характеристик, котрі дозволяють оцінити можливості універсальності:

- Ефективність – точність розпізнавання жестів на вхідному наборі даних.
- Швидкодія – можливість здійснювати переклад в реальному часі або кількість часу витрачена для перекладу жестів.
- Кросплатформеність – можливість використовувати програмні рішення на будь-яких платформах.
- Доступність – можливість користуватися програмними рішеннями при відсутності відповідних апаратних засобів.

Метою статті є проведення порівняльного аналізу існуючих систем сурдоперекладу для смартфонів, персональних комп'ютерів та окремих апаратно-технічних засобів, виділити ключові фактори для виявлення ряду переваг та недоліків кожного рішення з урахуванням вищезазначених критеріїв універсальності. Завдяки аналізу сформулювати кінцевий принцип системи сурдоперекладу з урахуванням алгоритму сурдоперекладу,

сформованого в статті «Аналіз сучасних систем розпізнавання дактильної-жестової мови для систем» [3] та можливостей кросплатформеності [4], а також стек технологій для створення відповідного програмного забезпечення.

Виклад основного матеріалу. В сучасному світі існують концептуально різні підходи в здійсненні комунікації глухонімих людей в суспільстві.

По-перше, це використання живого спеціаліста, сурдоперекладача, в реальному часі під час відеозв'язку. Такий принцип реалізований в сервісах типу Video Relay Service(VRS) [5] та Video Remote Interpreting (VRI) [6]. Сурдоперекладач підключається до відеоконференцій або відеодзвінків і надає переклад мовлення на мову жестів в умовах реального часу.

По-друге, це здійснення перекладу текстових повідомлень в різноманітних додатках. Подібний підхід реалізується в сервісах типу Text Relay Service (TRS) [7]. В користувачів є можливість комунікувати за допомогою текстових повідомлень, а сурдоперекладач надає переклад шляхом отримання повідомлень від сурдо користувача та передає їх іншому співрозмовнику.

По-третє, це використання вебплатформ або мобільних додатків в поєднанні із штучним інтелектом, котрі передаючи інформацію через текст, відео, аудіо чи жести. Подібні рішення є найбільш популярними і вони надають можливість наживо перекладати мовлення на мову жестів чи текст шляхом поєднання різних технологій.

Недоліками перших двох підходів, є наявність живого спеціаліста, адже дуже складно забезпечити принцип універсальності і підтримки 24/7 для всіх користувачів.

Як зазначалося в статті «Аналіз сучасних систем розпізнавання дактильної-жестової мови для систем» [3], кожен програмний чи апаратно-програмний спосіб сурдоперекладу повинен бути двонапрямний- комунікація до глухонімого і від глухонімого. Тому розглянемо найбільш популярні рішення, що допомагають глухим людям комунікувати у суспільстві з урахуванням вищезазначених критеріїв.

Для початку розглянемо додатки, що допомагають глухим людям комунікувати в суспільстві на прикладі Ava [8] або Google Live Transcribe [9] або Dragon Anywhere [10]. Такий тип додатків передбачає безпосередній функціонал спілкування за допомогою жестів у вигляді транскрипції мовлення в реальному часі, а також і набір додаткових опцій до наявного функціонала. Вони використовують технологію розпізнавання мовлення,

щоб перетворювати звукове мовлення на текст та відображають його на екрані пристрою, а також відтворювати за допомогою жестових мов. Недоліками подібних застосунків є відсутність деталізації мови жестів для конкретної країни чи регіону. Зазвичай є реалізація або найпопулярніших жестових мов, або реалізація найпопулярнішого набору жестів.

Подібний підхід реалізується і в додатках, котрі прямо не передбачають комунікацію за допомогою жестів. Існують застосунки для кастомізації екранних клавіатур смартфонів та планшетів – Gboard with ASL Stickers [11], Signily Keyboard [12]. Дані програми реалізують жестову мову у вигляді або наборів стікерів для жестової мови, які можуть бути використані під час текстової комунікації або як окрему мову на клавіатурі. Signily Keyboard використовує американську жестову мову (ASL) і містить жести та символи ASL [13]. Такі програми не дозволяють забезпечити для кожної жестової мови тип клавіатури або набір стікерів, бо розробники не зацікавлені у розв'язанні проблем глухонімих у світі, а керуються нормативно-правовими вимогами ринків чи регіонів в яких хочуть представити свій продукт.

Додатки, котрі дозволяють використовувати живих сурдоперекладачів по принципу VRS або VRI охоплюють найбільшу аудиторію в світі згідно з публічними даними. Наприклад, Sorenson Video Relay Service (SVRS) [14] або Be My Eyes [15] або RogerVoice [16] або Google Meet [17] або Microsoft Teams [18] та InterpreterNow [19] надають можливість сурдо споживачам здійснювати відеодзвінки через сурдоперекладача, який передає повідомлення у форматі мови жестів, але при цьому реалізуються абсолютно різний принцип приєднання спеціаліста до діалогу глухонімих людей. В таких програмах, як RogerVoice або Be My Eyes або Interpreter Now, передбачений функціонал для підключення оператора-сурдоперекладача за допомогою окремої кнопки/панелі кнопок, але кількість жестових мов, які охоплюють спеціалісти із підтримки додатків вкрай обмежена. В додатках Google Meet або Microsoft Teams наявний функціонал відображення кінцівок сурдоперекладача, але ніяких інструментів для автоматичного залучення перекладачів немає. Організатори відеодзвінків повинні самостійно знаходити та підключати відповідних спеціалістів.

Існують додатки або сервіси, що використовують набір жестів для конкретної жестової мови країни або містять набір різних мов для окремих країн – Spread Signs [20], Marlee Signs [21],

SignSchool [22], ASL Coach [23]. Слід зазначити, подібні додатки не мають прямого функціонала сурдоперекладу, а виступають лише як допоміжний інструмент – містять інтерактивні уроки та вправи або інші навчальні/тренувальні можливості в практикуванні жестової мови. Наприклад, Spread Signs [20] має великий словник жестів з можливістю пошуку жестів за словами або фразами. Marlee Signs [21] містить набір навчальних матеріалів і вправ для вивчення жестової мови, зокрема американської жестової мови – ASL [13]. Подібні рішення є гарним інструментом розширити навички сурдоперекладу, але ніяк не можуть виконувати самий сурдопереклад.

Найновіший спосіб сурдоперекладу реалізований в додатках, що використовують віртуального помічника або штучний інтелект – Hand Talk [21], Greta [22], SignAll Chat [23]. Реалізація перетворення тексту чи мовлення методами штучного інтелекту та віртуального помічника використовуючи різноманітні джерела зчитування інформації – камери, мікрофони чи дисплей. Такий тип додатків знаходиться на етапі розвитку і являти собою оптимальний, універсальний спосіб сурдоперекладу. За рахунок навчання штучного інтелекту можна збільшувати точність перекладу шляхом навчання на нових даних. Також віртуальний помічник – анімація людини дозволяє повністю прибрати людський фактор.

Проаналізувавши великий набір додатків для сурдоперекладу, їх переваги та недоліки, можна створити таблицю з порівнянням різних додатків сурдоперекладу.

Основними критеріями порівняння додатків в таблиці є апаратно-програмна відповідність, а саме якого типу рішення (мобільний додаток, розширення для додатку, програма для Unix чи Windows систем), платформи на які встановлюється (тип операційної системи на відповідній апаратній платформі).

Окремим критерієм порівняння є наявність кросплатформеності. Це дозволяє комплексно оцінити рішення з точки зору універсальності та можливості масштабування та адаптивності для різних країн та регіонів.

Для дослідження популярності візьмемо показники кількості завантажень та оцінку від користувачів із Google Play та AppStore. Слід зазначити, що в AppStore відображається лише оцінка від користувачів.

Останнім критерієм оцінювання застосунків є можливість підтримки та перекладу в реальному часі із графіком 24 години і 7 днів на тиждень.

Підтримка в реальному часі дозволить оцінити наскільки клієнтоорієнтований застосунок та як можна вирішувати нестандартні ситуації при користуванні додатком. Критерій «Підтримка перекладу людиною» дозволить оцінити наскільки ще актуальна проблема залежності сурдоперекладу від людини. Цей критерій напряду вже на проблеми сурдоперекладу і його залежності від людини.

В магазинах для додатків Google Play та AppStore пропонується велика кількість додатків, проаналізувати і дослідити, не мається можливості. Тому пропонується вибрати найпопулярніші рішення за кількістю завантажень та рейтингом, а також додати декілька специфічних додатків, що знайшлися в ході проведення дослідження.

Рішення	Тип рішення	Платформа	Крос платформеності	Кількість скачувань		Підтримка в реальному часі	Підтримка перекладу людиною
Ava	Mobile App	Windows/Mac/iOS/Android	+	GooglePlay	3.3 (100тис.+)	+	-
				AppStore	4.9		
Sorenson Video Relay Service	Services	-	+	-		+	+
Spread Signs	Mobile App	iOS/Android	-	GooglePlay	4.0 (500тис.+)	-	-
				AppStore	3.5		
Gboard	Mobile App	iOS/Android	-	GooglePlay	4.5 (5 млрд+)	-	-
				AppStore	3.7		
Marlee Signs	Mobile App	IOS	-	AppStore	3.3	-	-
SignSchool	Web/Mobile App	iOS/Android	-	GooglePlay	3.6 (100тис.+)	-	-
				AppStore	3.8		
Be My Eyes	Mobile App	iOS/Android	-	GooglePlay	4.4 (1 млн+)	+	+
				AppStore	4.9		
Signily Keyboard	Mobile App	iOS	-	AppStore	-	-	-
RogerVoice	Mobile App	iOS/Android	-	GooglePlay	3.9 (50тис.+)	-	+
				AppStore	4.7		
Google Meet	Web/PC/Mobil App	Windows/Linux/IOS/MAC/Android	+	GooglePlay	4.5 (5млрд+)	+	+
				AppStore	4.9		
Microsoft Teams	Web/PC/Mobil App	Windows/Linux/IOS/MAC/Android	+	GooglePlay	4.7(100 млн+)	+	+
				AppStore	4.6		
InterpreterNow	Desktop/Mobile App	Windows/iOS/Android	+	GooglePlay	4.2(10тис.+)	-	+
				AppStore	4.5		
Hand Talk	Mobile App	iOS/Android	-	GooglePlay	2.8 (1млн+)	-	-
				AppStore	4.0		
Greta	Mobile App	Android	-	GooglePlay	3.7(10тис.+)	-	-
SignAll Chat	Mobile App	iOS	-	AppStore	4.4	+	-
Google Live Transcribe	Extention/Mobile App	iOS/Android	-	GooglePlay	3.9(128тис.+)	-	-
				AppStore	4.6		
ASL Coach	Mobile App	iOS/Android	-	GooglePlay	2.6(1тис.+)	-	-
				AppStore	2.9		

Провівши аналіз таблиці з додатками для людей із вадами слуху та мовлення, виділяється певний набір додатків по популярності та технології перекладу.

Відповідно, Google Meet та Microsoft Team, є найпопулярнішими рішеннями, де до відео-конференції додається спеціаліст, коли потрібно

в реальному часі виконувати переклад жестів. Слід зазначити, що ці додатки не мають вбудованих, інтерактивних чи автоматизованих рішень для перекладу жестів без участі людини. Рішень від Gboard чи Signily Keyboard, скоріше, пропонують рішення для людей котрі вже володіють сурдоперекладом, адже при виборі мови жестів як

основного шару для екранної клавіатури смартфона, потрібно розуміти за що відповідає кожен із жестів. Такий підхід дуже сильно обмежує сферу застосування.

Окремим типом додатків, що мають відносно велику кількість звантаження (понад 2–3 тис.) це додатки в яких можна вивчати мову жестів. Наприклад, SignSchool або ASL Coach або Spread Signs спеціалізуються виключно на навчанні користувачів певних країн чи регіонів, де популярні відповідні жестові мови.

Спеціалізовані додатки, що розробляються виключно для перекладу жестів, на прикладі Marlee Signs або Google Live Transcribe або InterpreterNow або Ava, не є популярними та мають проблеми із підтримкою та подальшим розвитком та вдосконаленням. Слід зазначити, що Hand Talk або Greta або SignAll Chat використовують для перекладу засоби штучного інтелекту та комп'ютерного зору для розпізнавання, класифікації та перекладу жестів.

В результаті порівняльного аналізу можна стверджувати про наступні ключові речі:

– В сучасному світі відсутній єдине програмне або програмно-апаратні рішення для перекладу з чи на мову жестів. На цей момент, серед додатків популярним є використання живого сурдоперекладача і використання готової бази жестів і використання засобів штучного інтелекту.

– Відсутній єдиний підхід у розв'язання проблеми комунікації глухих та/або глухонімих людей у світі. ВООЗ та ООН лише говорять про проблему, але не про вирішення. Найбільші техно гіганти світі, Google або Microsoft, пропонують найпростіший підхід в сурдоперекладі – залучення живого перекладача.

– Відсутність обов'язкового двонаправленого режиму перекладу. Кожен із розглянутих додатків, що не використовує в якості перекладу живого спеціаліста, не має можливості перекладати з жестової мови і на жестову мову, тобто відсутність двонаправленого перекладу.

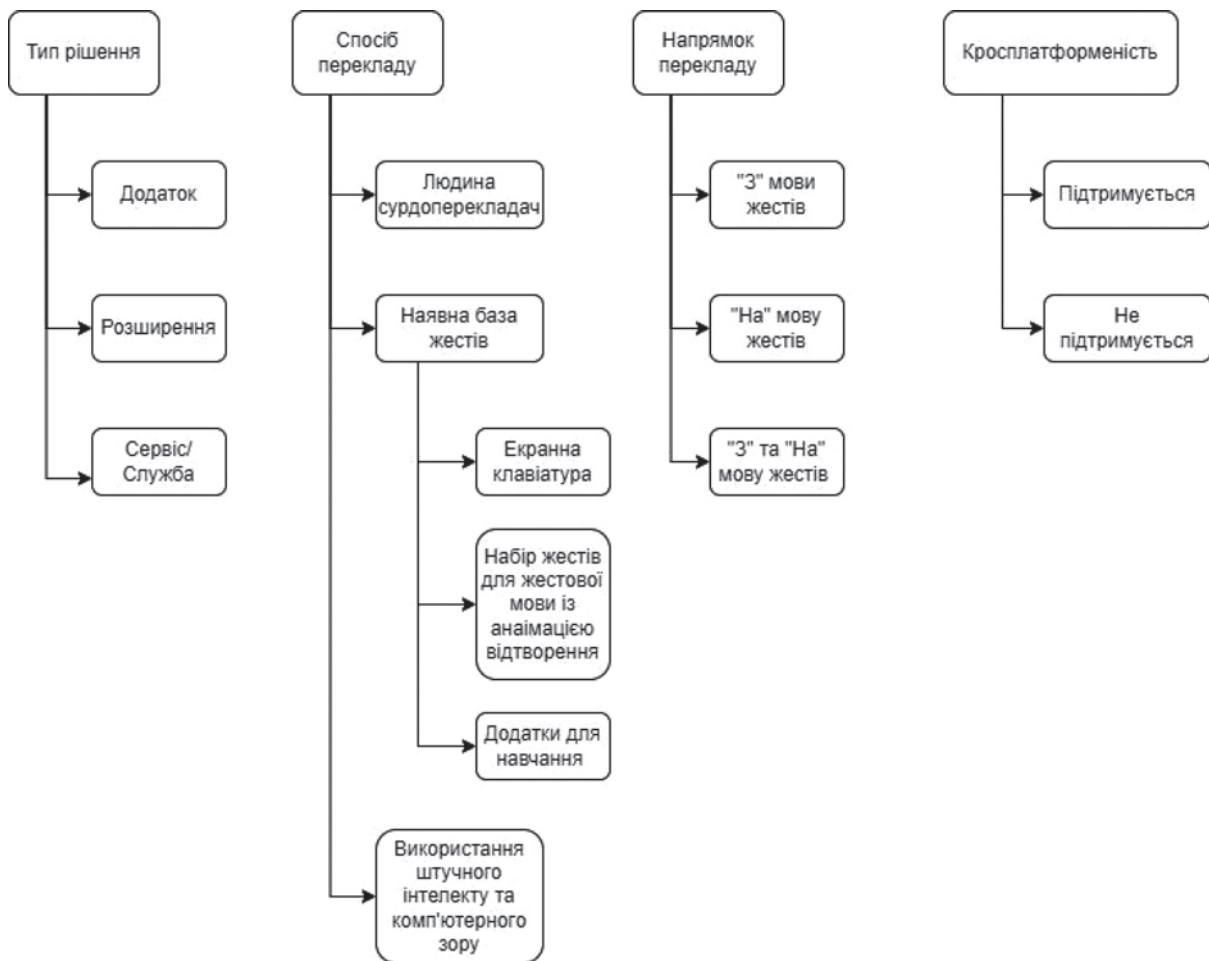


Рис. 1. Класифікація систем сурдоперекладу

Отже, на основі розглянутих підходів та рішень для сурдоперекладу, було розроблено нову класифікацію систем сурдоперекладу, котра представлена на рис. 1.

Висновки. Визначено, що існує багато систем сурдоперекладу для глухонімих людей, котрі спрощують момент спілкування в соціумі. Ці застосунки допомагають розпізнавати жести і перекладати на звичайну мову і навпаки, що в результаті прибирає бар'єр комунікації між глухонімими і здоровими людьми.

В результаті порівняльного аналізу з'ясовано, що застосунки в основі яких є використання живих спеціалістів мають найбільший попит. При цьому, слід зазначити, що такий тип застосунків поділя-

ється на наявність цілодобової підтримки (Be My Eyes, RogerVoice, InterpreterNow) та необхідності планувати дату та час підключення сурдоперекладача (SVRS, Google Meet, Microsoft Teams).

Різноманітність систем сурдоперекладу на базі смартфона чи персонального комп'ютера дозволяє структурувати всі рішення за рядом критеріїв, а саме типом платформи, підтримкою в реальному часі, можливістю кросплатформеності, популярності. Ця структура дозволяє оцінити більш глобально весь спектр рішення з наукової точки зору і на основі розширеної структурної класифікації запропонувати нові рішення. Також, даний аналіз буде корисний для споживачів, адже дозволяє зробити вибір на основі їх індивідуальних потреб та задач.

Список літератури:

1. WHO: 1 in 4 people projected to have hearing problems by 2050. URL: <https://www.who.int/news/item/02-03-2021-who-1-in-4-people-projected-to-have-hearing-problems-by-2050> (дата звернення 02.09.2023).
2. Launch of the World Report on Hearing. URL: <https://www.who.int/news-room/events/detail/2021/03/03/default-calendar/launch-of-the-world-report-on-hearing> (дата звернення 02.09.2023).
3. Лихошерстов Д.О., Лебедев Д.Ю. *Аналіз сучасних систем розпізнавання дактильної-жестової мови для систем сурдоперекладу*. Вчені записки ТНУ імені В.І. Вернадського. URL: <https://doi.org/10.3283/8/2663-5941/2021.6/07> (дата звернення 02.09.2023).
4. Лихошерстов Д.О., Лебедев Д.Ю. *Порівняльний аналіз шляхів визначення зображення на відео ряді засобами машинного навчання*. URL: <https://doi.org/10.32782/2663-5941/2022.5/03> (дата звернення 02.09.2023).
5. Video Relay Service (VRS). URL: <https://www.fcc.gov/vrs> (дата звернення 02.09.2023).
6. Video Remote Interpreting (VRI). URL: <https://www.washington.edu/doit/what-video-remote-interpreting> (дата звернення 02.09.2023).
7. Telecommunications Relay Services (TRS). URL: <https://www.fcc.gov/trs> (дата звернення 02.09.2023).
8. Professional & AI-Based Captions for Deaf & HoH | Ava. URL: <https://www.ava.me/> (дата звернення 02.09.2023).
9. Google Live Transcribe. URL: <https://transkriptor.com/> (дата звернення 02.09.2023).
10. Dragon Anywhere. URL: <https://www.nuance.com/dragon/dragon-anywhere.html> (дата звернення 02.09.2023).
11. Gboard (Google Keyboard) with ASL Stickers. URL: <https://uk.wikipedia.org/wiki/Gboard> (дата звернення 02.09.2023).
12. Signily Keyboard. URL: <https://aslized.org/innovations/signily/> (дата звернення 02.09.2023).
13. American Sign Language (ASL). URL: <https://www.nidcd.nih.gov/health/american-sign-language> (дата звернення 02.09.2023).
14. Sorenson Video Relay Service (SVRS). URL: <https://sorenson.com/vrs/> (дата звернення 02.09.2023).
15. Be My Eyes. URL: <https://www.bemyeyes.com/> (дата звернення 02.09.2023).
16. RogerVoice. URL: <https://rogervoice.com/en/> (дата звернення 02.09.2023).
17. Google Meet – Online Video Calls, Meetings and Conferencing. URL: <https://meet.google.com/> (дата звернення 02.09.2023).
18. Microsoft Teams. URL: <https://www.microsoft.com/uk-ua/microsoft-teams/log-in/> (дата звернення 02.09.2023).
19. InterpreterNow. URL: <https://signvideo.co.uk/> (дата звернення 02.09.2023).
20. Spreadthesign. URL: <https://www.spreadthesign.com/uk.ua/search/> (дата звернення 02.09.2023).
21. Marlee Signs. URL: <https://www.marleematlin.net/app/> (дата звернення 02.09.2023).
22. SignSchool: Learn ASL for Free. URL: <https://www.signschool.com/> (дата звернення 02.09.2023).
23. ASL Coach. URL: <https://www.youtube.com/watch?v=7n-K0WmjigM> (дата звернення 02.09.2023).
24. Hand Talk. URL: <https://www.handtalk.me/en/app/> (дата звернення 02.09.2023).
25. Greta application. URL: <https://play.google.com/store/apps/details?id=com.pack1.deafcommunication&hl=en> (дата звернення 02.09.2023).

Lykshosherstov D.O., Lebedev D.Yu. COMPARATIVE ANALYSIS OF DIGITAL SIGN LANGUAGE INTERPRETATION TOOLS

In recent years, there has been significant progress in the development, adaptation, and implementation of sign language interpretation systems in society to simplify the lives of people with hearing and speech impairments.

The article provides a comparative analysis of sign language interpretation systems aimed at identifying the advantages and disadvantages of each system for improving communication for people with hearing impairments. An analysis of current research and approaches presented in various applications, plug-ins, or extensions for applications regarding the functionality of current sign language interpretation systems based on smartphones, personal computers, and individual hardware and software solutions was performed. The research analyzed several solutions that use conceptually different approaches to sign language interpretation. The most popular approach is the use of live sign language interpreters, who can be either professional specialists or volunteers. A relevant solution is to use ready-made sign databases for the most popular sign languages, where it is possible to provide translation for inexperienced users. The most modern approach is to use video image recognition and artificial intelligence technologies.

Unfortunately, the actual number of unique solutions and, accordingly, software implementations in the form of applications do not allow us to fully assess the problems of sign language interpretation, but according to current research, this is not the main criterion of the study.

Key words: *sign language interpretation, gestures, deafblind, smartphone, application, artificial intelligence, comparative analysis.*