

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Київський національний університет імені Тараса Шевченка
Інститут філології
Кафедра української мови та прикладної лінгвістики

**СТВОРЕННЯ АВТОМАТИЧНОГО ГЕНЕРАТОРА
ФОНЕТИЧНОЇ ТРАНСКРИПЦІЇ ДЛЯ
УКРАЇНСЬКОМОВНИХ ТЕКСТІВ**

Кваліфікаційна робота

освітнього ступеня «бакалавр»
за спеціальністю 035 «Філологія»,
спеціалізацією 035.10 «Прикладна
лінгвістика»,
галузі знань 03 «гуманітарні науки»
ОПП «Прикладна (комп'ютерна)
лінгвістика та англійська мова»

Яни ГРИЦЕНКО

Наукові керівники:

к. філол. н., доц. Оксана ЗУБАНЬ
Валентина РОБЕЙКО

Рецензент:

канд. техн. наук Микола САЖОК

ЗМІСТ

ВСТУП	3
РОЗДІЛ 1. Теоретико-методологічна парадигма досліджень українських фонетичних транскрипцій	5
1.1. Історія та розвиток представлення звуків мовлення у вигляді транскрипції. Види транскрипцій	5
1.2. Сучасний стан досліджень української фонетичної транскрипції: актуальні питання та проблеми	11
Висновки до розділу 1	14
РОЗДІЛ 2. Використання автоматичної фонетичної транскрипції в технологіях опрацювання та синтезу мовлення	15
Висновки до розділу 2	22
РОЗДІЛ 3. Створення автоматичного генератора фонетичної транскрипції для українськомовних текстів	23
3.1. Розробка транскриптора для українськомовних текстів	23
3.2. Правила транскрибування	31
3.3. Тестування розробленої програми транскрибування текстів	37
3.4. Плани та перспективи покращення автоматичного транскриптора	42
Висновки до розділу 3	43
ВИСНОВКИ	44
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	46
ДОДАТКИ	50

ВСТУП

Актуальність теми нашого дослідження зумовлена активним розвитком технологій обробки природної мови та їх популярності на сьогоднішньому ринку. Синтезатори знаходять багато практичних застосувань у сучасному світі: завдяки ним уможлиблюється вивчення нових іноземних слів з правильною вимовою, прослуховування аудіокниг, створених без залучення живого диктора, озвучування тексту презентації реалістичними голосами тощо. На сьогодні користувачеві представлено вибір із безлічі програм, доступних для використання як на комп'ютерах, так і на переносних гаджетах. На сучасному етапі розвитку українських синтезаторів мовлення є необхідним створення фонетичної транскрипції тексту для блоку лінгвістичної обробки. Актуальність дослідження зумовлюється також важливістю опрацювання тексту на фонетичному рівні разом із морфологічним, синтаксичним та семантичним аналізами тексту у системах обробки мови.

Метою дослідження є створення програми для автоматичної генерації фонетичної транскрипції українською мовою, яка може бути використана у технологіях обробки тексту та синтезу мовлення.

Реалізація поставленої мети передбачає вирішення таких **завдань**:

- дослідження фонетичної системи української мови, здійснення аналізу інтерпретацій поняття фонем та алофону та змін, що відбуваються зі звуками у мовленнєвому потоці;
- дослідження видів транскрипцій та способів позначення змін звуків української мови у мовленнєвому потоці за допомогою фонетичного алфавіту у транскрибованому тексті;
- аналіз наукових праць про сучасні системи обробки та синтезу тексту, визначення ролі фонетичної транскрипції у лінгвістичному блоці синтезаторів штучного мовлення;

- розробка алгоритму для транскрибування тексту українською мовою та реалізація програми в інтегрованому середовищі для програмування мовою Python;
- тестування транскриптора на текстових масивах та виправлення помилок у його роботі.

Об'єктом дослідження є звуковий і фонологічний склад української мови. **Предмет дослідження** – графічне відображення звуків української мови засобами фонетичного алфавіту. **Матеріал дослідження** – українськомовні літературні тексти та комп'ютерні інструментальні засоби.

Поставлені завдання були реалізовані за допомогою таких **методів**: загальнонаукові (опис, аналіз та синтез), теоретичні (теоретичний аналіз, дефінітивний аналіз), порівняльно-історичний.

Наукова новизна роботи полягає у розробці програмного модуля, що здійснює відображення звукового складу тексту українською мовою у вигляді фонетичної та фонематичної транскрипції. Розробка може бути використана як в теоретичних, так і в практичних цілях для дослідження звукового та фонематичного складу української мови, застосування у навчальному процесі, зокрема при вивченні української мови іноземцями, та можливого інтегрування до систем синтезу штучного мовлення.

Структура роботи: Робота складається зі вступу, двох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг роботи складає – 54 стр.

РОЗДІЛ 1. Теоретико-методологічна парадигма досліджень українських фонетичних транскрипцій

У цьому розділі розглянемо погляди різних дослідників на транскрипцію та її використання протягом історії її розвитку, порівняємо визначення поняття фонемі й алофону та дослідимо внесення українських мовознавців у систему символів для передачі звуків мовлення при транскрибуванні. Також у даному розділі ми представимо сучасні питання та проблеми української транскрипції, підняті у працях різних науковців.

1.1. Історія та розвиток представлення звуків мовлення у вигляді транскрипції. Види транскрипцій

Транскрипцією називають графічний запис живого мовлення з використанням спеціально розробленого фонетичного алфавіту. Дослідник А. Соколянський пропонував розвиток видів фонетичних транскрипції як один із напрямів практичного використання багаторівневої філології.

Говорячи про представлення звуків мовлення у вигляді фонетичної транскрипції, вважаємо доцільним розглянути поняття звуку та фонемі. Уперше термін «фонема» був використаний у 1873 році французьким лінгвістом А. Дюфріш-Дженетом, пізніше термін запозичив Ф. де Сосюр, який вважав фонему «одиницею системи вираження звукової мови». І. Бодуен де Куртене наголосив на різниці між «звуком мовлення» та фонемою, назвавши звук фізично-акустичним явищем, а фонему – його психічним еквівалентом. О. Курило давала визначення фонемі як психофонетичній одиниці, у якій поєднані об'єднані всі звукові відтінки, Л. В. Щерба називав фонемами звукові типи, у які об'єднано звуки мовлення, що здатні «диференціювати слова та їх форми» [29], М. Наконечний визначав фонему як найменшу звукову одиницю, яка знаходить конкретний вияв у варіантах. Тлумачення терміну «фонема» О. Синявського підтримує тлумачення Щерби, визначаючи фонему як не просто один із безлічі звуків людської мови, а таку мінімальну звукову

одиницю в звуковому процесі, звукотип, яка спроможна ввести розрізнення до значення слова.

При цьому, на думку О. Синявського, кількість таких одиниць в мові є обмеженою. За словами вченого, окремий звук тільки тоді може бути визнаний фонемою в певному слові, якщо в мові існує слово, значення якого відрізняється від даного лише цією фонемою, а усі інші фонемі є спільними. У той час ті звуки, які не диференціюють значення слів, є лише відтінками фонем, тобто алофонами, як, наприклад, «білабіальне та лабіадентальне [в]». На думку Н. Тоцької, фонема є саме узагальненням тих чи інших звучань – алофонів, що властиві тільки їй. Кількість фонем, утворених алофонами, яких у мові існує безліч унаслідок потрапляння фонемі до різних фонетичних умов, є обмеженою.

Сама Н. Тоцька виділяє 6 голосних фонем і 32 приголосні фонемі в звуковій системі української мови. У схемі мовного знака вона вважає фонему компонентом, що виконує функцію позначення, а тому слугує для забезпечення процесу комунікації. У сучасній українській лінгвістиці широкоприйнятим є її визначення фонемі як «мінімальної одиниці звукової будови мови, функція якої полягає у розрізненні і розпізнаванні значущих одиниць — морфем, а через них і слів» [29, 22].

Історія розвитку транскрибування звуків почалася ще з XVI ст., коли були відомі різноманітні системи фонетичних транскрипцій, запропоновані такими вченими як Мегре (Франція), Буллокар, Гарт, Сміт (Англія). Власні системи, розроблені багатьма науковцями, найчастіше зводилися до таких способів звукового запису: ті, що базувалися на використанні латинського алфавіту та позначали звукові варіанти діакритичними знаками та видозміненими літерами й ті, що користувалися абсолютно новими створеними знаками. II половина XIX століття – I половина XX характеризуються появою ряду нових систем фонетичних транскрипцій, розроблених науковцями. Так, К.Р. Лепсіус, німецький історик-археолог та дослідник древніх текстів, запропонував позначати такі відтінки звуків як

лабіалізація, пом'якшення, довгота, нескладотворчість голосних, какумінальність, апікальність приголосних, носове забарвлення і т.п. різноманітними діакритичними знаками у своїй системі транскрипції.

Варто також зазначити, що 1888 року міжнародною фонетичною асоціацією було розроблено найбільш широко відомий та уживаний принцип фонетичного позначення, що відомий під назвою Міжнародний фонетичний алфавіт (МФА). Спершу система була створена для транскрибування французького мовлення за допомогою латинського алфавіту. У 1925, 1947, 1951 роках алфавіт доповнився великою кількістю діакритичних знаків та альтернативними позначеннями літер, що дало змогу надати системі деякої універсальності. Над підвищенням універсального характеру системи працювали також такі мовознавці, як Г. Сміт, О. Єсперсен та ін. Метою нововведень було уможливлення передачі якомога більшої кількості існуючих звуків мов світу. Вклад Г. Сміта був вагомим: фонетист увів до застосування 109 нових позначень. Проте й цього не було достатньо для відтворення в записі усіх відтінків звуків, що зустрічаються у мовах різних народів.

Оригінальна так звана анальфabetична система звукозапису була розроблена датським лінгвістом Отто Єсперсеном. В її основі лежало вираження звука за допомогою символічної формули, яка використовувала цифри й букви латинського та грецького алфавітів для описання способу й місця артикуляції того чи іншого звука. «Грецькою літерою позначається рухомий орган артикуляції, латинською – нерухомий (тобто місце артикуляції), а цифрою – форма отвору для проходу струменя повітря...». Закцентуємо на тому, що дана система була використана в українському мовознавстві лінгвістом І. Зилинським. І. Зилинський доповнив та уточнив метод, запропонований О. Єсперсом, описуючи приголосні звуки української мови.

Що стосується російської лінгвістики, натепер у ній теж досі не сформувалася єдина загальноприйнята система фонетичної транскрипції.

Системи з тою чи іншою кількістю альтернативних позначень літер та діакритичних знаків, в основі яких лежить латинська чи російська графіка, використовуються різними розробниками. Л. Г. Зіндер викладає такі причини варіативності способів звукового запису, як відсутність розмежування між фонетичною та фонематичною транскрипціями, що є двома основними видами звукового запису; відсутність такої класифікації мовних звуків, що лежать в основі транскрипції, яка була б загальноприйнятою, та інертність тієї звички, що склалася у результаті впливу шкіл і традицій [31].

Власне, в українському мовознавстві також присутня різноманітність використовуваних систем фонетичної транскрипції. У своїх доробках такі дослідники як І. Зилинський та О. Брок доповнювали транскрипцію, в основі якої лежало використання латинської графіки та діакритичних знаків, буквами грецького та українського алфавітів. М. Ф. Наконечний розробив докладну систему фонетичної транскрипції, взявши за основу український алфавіт та доповнивши його, подібно до своїх попередників, літерами латинського та грецького алфавітів, а також діакритичними знаками та надрядковими символами. Пізніше цю систему було включено до розділу «Фонетика» в академічному «Курсі сучасної української літературної мови» (ред. Л. А. Булаховський). Відмітимо також систему М. Ф. Наконечного, яка внесла вагомий вклад у розвиток українського мовознавства. Так, багато українських науковців, серед яких автори підручників М. А. Жовтобрюх та М. П. Івченко, користуючись транскрипцією М. Наконечного як основою, здійснили розробку власних систем.

Маючи на меті позбутися такої значної кількості різноманітних систем фонетичної транскрипції, 1959 року при інституті мовознавства АН СРСР було скликано спеціальну комісію для розробки уніфікованої системи транскрипції, презентованої у двох варіантах: на основі російського алфавіту та на основі латинського алфавіту. Таким чином було запропоновано декілька проєктів для мов різної структури, представлених радянськими народами. Важливо зазначити, що 1962 року спеціальною Республіканською

комісією при Інституті мовознавства АН УРСР було запропоновано ряд принципів та рекомендацій їх реалізації, на основі яких мають базуватися фонетична та фонематична транскрипції української мови.

Важливим вважаємо зазначити, що фонематична транскрипція також пройшла свої етапи розвитку. У XIX в науковому просторі відбувся розвиток фонології як «науки про функціонування одиниці звукової системи мови», і разом з тим виникла потреба у такому виді наукової транскрипції, який не брав би до уваги позиційні зміни звуків та алофони, позначаючи лише фонемний склад морфем та слів, на відміну від фонетичної транскрипції, метою якої є передання всіх можливих відтінків звуків у мовленні. І. Бодуен де Куртене описував важливість використання окремих транскрипцій для фонем і їх варіантів. Значною роботою в цій сфері стали розроблені англійським мовознавцем Г. Світом дві системи транскрипції: перша, так звана «широка», система, була створена для позначення «ширших смислорозрізнявальних звуків», що виконують дистинктивну функцію; друга, так звана «вужька», транскрипція, мала на меті відображення відтінків звуків при вимові.

Л. В. Щерба у своїх наукових роботах наголошував на важливості протиставлення фонетичної та фонематичної транскрипцій: «Запис мовлення за допомогою фонетичного алфавіту на основі суворо проведеного звукового принципу називається транскрипцією. При цьому, якщо ми будемо записувати тільки фонемі..., то одержимо фонологічну, або фонематичну транскрипцію. Коли ж ми будемо записувати й усі відтінки, в яких фонема існує в мові, то одержимо фонетичну транскрипцію» [32].

Більш того, на думку деяких мовознавців, варто виділяти навіть три різновиди науково-лінгвістичних транскрипцій. Наприклад, Р. І. Аванесов запропонував такі їх види: фонетична, словофонематична та морфофонематична. Як бачимо, фонематичну транскрипцію мовознавець вважає за доцільне поділити на два вужчі типи для різного використання. Завдання трьох запропонованих транскрипцій Р. Аванесов викладає

наступним чином. Фонетичну транскрипцію дослідник вважає емпіричною, тобто такою, що цілком не має зв'язку зі структурою самої мови, а натомість лише передає вимову як таку [1, 216]. При цьому фонетична транскрипція, як стверджує вчений, має позначати найкоротші одиниці мови в «усьому багатстві їх фізіолого-акустичних властивостей». Словофонематична транскрипція розглядає слова тільки в певній формі, залишає лише «внутрішнє», самостійне й функціонально значуще звукової одиниці, позбавляючи звуки всього «зовнішнього», що зумовлене їх позиційним оточенням; таким чином словофонематична транскрипція «передає звукову оболонку конкретного мовного факту» [32]. Нарешті, морфофонематичний запис звуків передає фонетичну структуру уже не слів, а окремих морфем, розглядаючи їх тільки в позиції максимальної диференціації, тобто в такій, де проявляються функціонально значущі їх сторони. Досягається це, знову ж, звільненням звукової одиниці від будь-яких позиційних впливів. Зазначимо, що розробки Р. Аванесова критикуються в науковому оточенні за складність та деяку непослідовність.

Ми підтримуємо думку українських дослідників, які вважають, що в українській науці необхідним є застосування саме двох основних видів науково-лінгвістичної транскрипції: фонетичного та фонематичного. Обґрунтування їх використання та диференціація наявні у багатьох підручниках загального мовознавства. Так, в літературі української лінгвістики принципи фонематичного запису слів були вперше сформульовані О. Н. Синявським у праці, в якій автор ставив за мету дослідити частоту вживання фонем у творах українських письменників та в українській пресі, попередньо встановивши фонетичну систему сучасної української мови. О. Синявський зазначає, що послуговувався цілком транскрибування вибраних творів таким чином, аби означити «фонемі і лише фонемі», тобто ті чи інші звукові відтінки мали бути позначені як їх найближчі фонемі. Транскрипцію О. Синявського, проте, можна назвати скоріше неповною фонетичною, ніж фонематичною. Причину тому можна

знайти в своєрідному розумінні автором поняття фонемі [25]. У цілому в українському мовознавстві питання виділення фонематичної (або як її ще називають - фонологічної) транскрипції виникло тоді, коли з'явилася потреба розмежувати фонемі як одиниці системи фонем сучасної української літературної мови та їх варіанти як різні відтінки реалізацій звуків; характеризувати фонетичну систему української мови.

1.2. Сучасний стан досліджень української фонетичної транскрипції: актуальні питання та проблеми

Дослідженням питань фонетичного відображення вимови слів в українській мові займаються і сучасні дослідники. Зокрема, інтерес представляє передача транскрипцією слів іншомовного походження. Так, наприклад, у своєму дослідженні аббревіатур Р. Микульчик піднімає питання запису іншомовних аббревіатур в українській транскрипції, зазначаючи, що вони не мають стандартизованого написання [18]. У 2000 році Н. Кінець та М. Кратко розглянули «проблеми транскрипції в укладанні англійсько-українських словників з обчислювальної техніки, математики й програмування». Дослідники зазначають, що питання однозначності транскрибування термінології є важливим зокрема для точних та комп'ютерних наук, що зумовлено обробкою тексту обчислювальними машинами та необхідністю ідентифікувати однакові поняття. На думку дослідників, запис іншомовних слів із використанням українського алфавіту, що враховує вимову, не є абсолютно вдалим. Наприклад, у науковій літературі англійська фонема /æ/ у власних назвах часто передається то за допомогою [e], то за допомогою [a]. Н. Кінець та М. Кратко пропонують користуватися позначеннями із словника «Everyday's English Pronouncing Dictionary» Д. Джонса у якості еталону знаків транскрипції для передачі звукового складу запозичених з англійської мови слів [12].

Інтерес для сучасних дослідників представляє також практична транскрипція іншомовних слів. Лінгвісти А. А. Реформатский [22],

Ю.С. Маслов [17] та П. Кабберлі [8] окрім фонетичної та фонематичної виділяли ще й «практичну» транскрипцію. Використовувати її пропонували за необхідності введення слів чи словосполучень до складу іншої мови [19]. Передача звукового складу слова має відбуватися за допомогою реальних літер алфавіту даної мови без використання діакритичних знаків та з урахуванням звичної орфографічної системи. Така система транскрибування піддається критиці. У випадку з використанням латиниці проблема виникає у розбіжності складу алфавітів та орфографічних правил у мовах різних народів світу [19]. Практичне транскрибування українською мовою піддавалося критиці М.О. Вакуленком, не Проблему дослідник бачив у поєднанні в такій транскрипції транслітерації й граматичного транскрибування, що призводило до одночасного транслітерування та транскрибування різних букв та звуків [27]. Сам М. Вакуленко розрізняв ще й такі види транскрипції як інваріантна та граматична [28, 7]. Вагомий вклад у розвиток фонетики сучасної української мови внесли Н. І. Тоцька [28], М. А. Жовтобрюх [7], О. В. Бас-Кононенко [3].

Як уже зазначалося вище, для наукового фонетичного запису необхідним є створення фонетичного алфавіту – системи літер із використанням додаткових діакритичних знаків, що слугують для точного відтворення звукової системи мови. Найчастішим є використання латинської графіки, але існують фонетичні алфавіти на основі азбуки окремих мов чи, як ми дослідили в цьому розділі, їх поєднання. В українській лінгвістиці використовують транскрипцію, розроблену на базі українського алфавіту з використанням таких латинських знаків: різні варіанти фонем */v/* - губно-губний та губно-зубний – відповідно позначають символами [w] та [v]; для позначення ж середньоязикового щілинного сонорного приголосного, що в позиції перед голосними виступає як варіант фонем */j/*, використовують [j], як у словах: [jakip], [jomu], [juна] і т.д. [31].

Український фонетичний алфавіт відповідає вимогам, що ставляться до фонетичного алфавіту Міжнародною Фонетичною Асоціацією (МФА): 1) кожен звук повинен мати спеціальний знак, який відповідає йому на письмі; 2) кожен знак має позначати лише один звук, тоюто бути вживаним тільки в одному значенні; 3) система має мати аналітичний характер, тобто спільну для ряду звуків додаткову ознаку потрібно могли позначити спільним додатковим знаком (знак лабіалізації для позначення огублення, хвилька для позначення назалізації голосного звука тощо); 4) система має бути зрозумілою та зручною для використання.

У 2010-х роках О. Стеріополо написала статтю, у якій дослідила українську фонетичну систему в парадигмі МФА. Розглядаючи фонетичні характеристики звуків фонетичної мови, дослідниця зазначила, що транскрипції, які використовують символи міжнародного фонетичного алфавіту, – як фонетичні, так і фонематичні – пропонують лише відносне уявлення про фонетично-фонологічні системи мов світу, і не здатні надати його суцільно точний опис [27, 192]. О. Стеріополо здійснила спробу порівняння дійсних традицій опису приголосних та голосних у системі української мови, зокрема уточнення місця артикуляції приголосних та місце /a/ й /и/ на трапеції голосних. Так, наприклад, у таблиці приголосних української мови МФА на противагу «пом'якшеним алофонічним варіантам» [dʒ, tʃ, ʒ, ʃ] африкат /dʒ, tʃ/ і фрикативних /ʒ, ʃ/, під час артикуляції котрих відбувається просунення язика у напрямку нижніх альвеол, символами [dʒ, tʃ, ʒ, ʃ] відображено їх «ретрофлексну реалізацію» Що стосується голосних, то, за словами дослідниці, українська фонема /и/, що в МФА позначається як /i/, має своєрідний характер, оскільки місце творення цього звука має суттєві відмінності від положення його позначення в трапеції голосних, розробленій МФА. Український голосний /и/ вимовляється у передній частині мовного апарату з деяким зміщенням назад у ротовій порожнині. Це дозволяє класифікувати його як голосний «переднього ряду зниженого підняття» [26, 52]. На трапеції ж голосних /и/ розміщено у «середньому ряді верхнього

підняття», що є притаманним, для прикладу, російській мові. У той же час наголошені голосні [ε] та [ɔ] наділені «більш відкритим характером» за [e] та [o], що ними транскрибують /e/ й /o/ в українському мовознавстві. Стеріополо продемонструвала місце українських голосних у позначенні символами МФА на трапеції голосних (Рис 1).

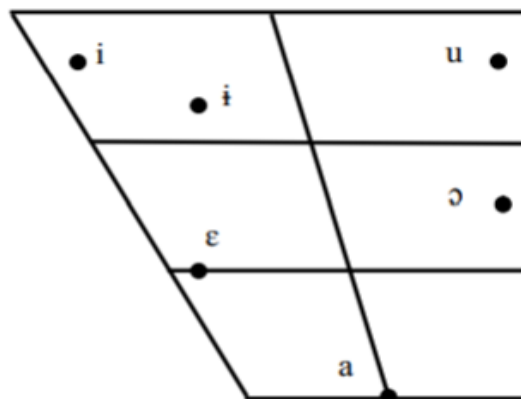


Рис 1. О. Стеріополо.

Наголошені голосні української мови [26, 55]

Висновки до розділу 1

Таким чином, сьогодні в українському мовознавстві практичне використання знаходять два види транскрипції: фонетична та фонематична. Першою послуговуються за необхідності як можна точніше передати всі відтінки звуків мовлення, а другою – за потреби дослідити систему фонем української мови. Питаннями коректного представлення звукової системи української мови в МФА (міжнародному фонетичному алфавіті) займаються сучасні дослідники: з'ясовано, що позначення деяких варіантів фонем є точнішим, ніж за допомогою фонетичного алфавіту української мови, у той час як місце деяких голосних на трапеції голосних звуків репрезентовано не точно. Нерозв'язаним у сучасній фонетиці залишається питання усталеного транскрибування слів та термінів іншомовного походження.

РОЗДІЛ 2. Використання автоматичної фонетичної транскрипції в технологіях опрацювання та синтезу мовлення

У другому розділі ми розглянемо переваги створення автоматичних транскрипцій та їх використання у сучасній практичній лінгвістиці, здійснимо аналіз розвитку технологій розпізнавання та синтезу мовлення та з'ясуємо необхідність автоматизації процесу транскрибування в лінгвістичних блоках TTS.

Технології опрацювання живого мовлення активно розвиваються протягом останнього століття, і одним із передових завдань ІТ-компаній є розробка синтезаторів Text-to-Speech (TTS), що забезпечують синтез мовлення, наприклад, для голосових помічників та чат-ботів. Говорячи про практичне застосування фонетичної транскрипції, варто розглянути такий пласт інформаційно-комунікаційних технологій як аналіз та синтез мовлення, а також описати місце, яке в процесі їх створення займає подання тексту в вигляді транскрипції [37].

Поява комп'ютерів і їх проникнення в різноманітні сфери людського життя привели до створення і розвитку особливих напрямів в комп'ютерних технологіях, які пов'язані зі звуковою мовою. Ці напрямки відповідно до власних потреб отримали назви "автоматичний синтез мовлення" і "автоматичне розпізнавання мовлення". За останні 10-15 років відбулися помітні зміни в комп'ютерних технологіях синтезу і розпізнавання мовлення, обумовлені постійно зростаючими можливостями комп'ютерної техніки, вдосконаленням математичного апарату і більш глибоким розумінням реальних процесів породження і розуміння мови людиною. Зазначимо, що сьогодні використання комп'ютерів у прикладній фонетиці є широко поширеним: трудомісткі завдання, які раніше виконувались вручну, тепер можуть виконуватися автоматично, забезпечуючи меншу витрату часу, фінансів, а часто й більшу точність та якість. Одним із таких особливо трудомістких, витратних та схильних до помилок завдань є фонетична

транскрипція [37, 3]. Справедливим буде стверджувати, що одним із перших важливих кроків використання інформаційних технологій в лінгвістиці є дигіталізація тексту – переведення мовного матеріалу, що існує у друкованому та усному виді, в цифрову форму [40].

Розпізнавання мовлення, також відоме як автоматичне розпізнавання мовлення (АРМ), комп'ютерне розпізнавання мовлення або перетворення мовлення в текст, – це можливість, яка дозволяє програмі обробляти людську мову в письмовому форматі. Хоча його зазвичай плутають із розпізнаванням голосу, розпізнавання мовлення зосереджується на перекладі мови із словесного формату в текстовий, тоді як розпізнавання голосу просто намагається ідентифікувати голос окремого користувача. Автоматичний аналіз та синтез тексту є важливими задачами комп'ютерної лінгвістики як з точки зору розвитку теорії (розробки лінгвістичних основ створення штучного інтелекту), так і з точки зору реалізації практичних потреб людини, наприклад, створення голосових помічників.

Сучасний синтез мовлення є продуктом багатої історії спроб генерування мовлення механічними засобами. Найдавніший відомий пристрій для імітації людського мовлення був сконструйований Вольфгангом фон Кемпеленом понад 200 років тому. Його машина складалася з елементів, що імітували різні органи, що використовуються людьми для виробництва мови – міх для легенів, трубка для голосового тракту, бічна гілка для ніздрів тощо. Інтерес до таких механічних аналогів голосового апарату людини продовжувався і надалі. У другій половині XIX століття Гельмгольц та інші вчені почали синтезувати голосні та інші соноранти шляхом накладання гармонійних форм хвиль з відповідною амплітудою. Значно інший напрямок було прийнято Дадлі в 30-х роках минулого століття, коли він відкрив природу носія мови та її наслідки – так звану «модель джерела-тракту». За допомогою клавіатури для управління змінним у часі фільтром він зміг синтезувати досить якісну, вільну мову [9].

У середині 60-х років в розробках автоматичного синтезу мови сталося важливе зрушення: замість ресинтезу, або відновлення, обмеженої кількості мовних зразків, що зберігаються в комп'ютері в цифровій формі, була поставлена задача озвучування будь-якого повідомлення, що вводиться в комп'ютер в текстовому вигляді і невідомого заздалегідь системі звукового синтезу. Це призвело до розвитку синтезаторів типу "Текст-Мовлення" (Text-to-Speech або скорочено TTS). В ідеалі такі пристрої повинні імітувати діяльність людини, яка читає письмове повідомлення або текст будь-якого ступеня складності.

Перший цифровий синтезатор мовлення був продемонстрований в лабораторіях Белла приблизно в 1967 році. У деякому сенсі цей синтезатор був поверненням до машини фон Кемпелена, з однією головною відмінністю. Замість того, щоб маніпулювати механічною моделлю, синтезатор обчислював, що зробить механічна модель, якби вона була реалізована. На основі правил, отриманих під час вивчення людського мовлення, комп'ютерна програма обчислює послідовність форм, через які повинен пройти голосовий тракт, щоб генерувати мовлення, що відповідає будь-якому тексту, представленому як вхідні дані. На основі цих форм та знань про відповідне акустичне збудження (періодичні імпульси для дзвінких звуків, шумоподібне збудження для неголосих звуків) програма «вирішила хвильове рівняння з відповідними граничними умовами для обчислення акустичного тиску на губи» [9]. Нарешті, до гучномовця подавався електричний сигнал з такою ж формою хвилі, що і обчислюваний тиск, для отримання бажаної мови.

Цей метод синтезу мовлення має сильну привабливість, оскільки імітує спосіб, яким людина виробляє мовлення. Однак, незважаючи на значні зусилля, ще не доведено, що можна автоматично генерувати якісну мову із довільного тексту цим методом. Це пов'язано з тим, що не вдалося вивести правила, які працюють належним чином за будь-яких обставин. Поки що хороший синтез за правилами, як вважається, вимагає занадто частої

спеціальної модифікації правил, щоб бути корисним на практиці. Крім того, важко генерувати певні мовні звуки (наприклад, сплески в звуках, таких як /к/ кошеня) за правилом [40].

Один із різновидів сучасного синтезу тексту в мовлення (TTS) базується на набагато менш фундаментальній, але набагато ефективнішій процедурі, яка називається конкатенативним синтезом. В основному, бажаний мовний сигнал збирається з “одиниць”, вибраних з інвентаризації, складеної під час навчальної фази синтезатора. Ці одиниці – це акустичні зображення невеликих елементів підслів (наприклад, фонем, дифони, часто зустрічаються кластери приголосних тощо). Саме подання може мати декілька альтернативних форм. Повний процес конкатенативного синтезу, від введення тексту до виводу мови, складається з декількох етапів [40]. Б.М. Лобанов описав принцип роботи одного з таких синтезаторів російського мовлення “Фонемафон-2000” [35]. Так, після попередньої обробки тексту, що включає в себе такий важливий етап як уведення правил вербалізації чисел, скорочень та аббревіатур [16], [39], відбувається його перетворення на розмічений фонемний текст, який розбивається на фонетичні періоди, фрази та синтагми; процес просодичної розмітки відбувається у блоці пофразової обробки тексту. Далі текст передається до блоку послівної обробки, який звертається окремо до кожного слова, а не до цілої фрази; тут проставляються наголоси, відбувається об’єднання фонетичних слів та здійснюється фонемне транскрибування. Наступним етапом структури роботи синтезатора є власне просодичний процесор. Тут визначається інтонаційний тип синтагм, відбувається їх розбиття на акцентні групи та розбиття акцентних груп на елементи акцентних груп, і, нарешті, формування просодичних параметрів із використанням еталонних портретів інтоном.

Останні сучасні нейронні моделі синтезу тексту в мовлення (TTS) значно покращили зрозумілість і природність генерованого мовлення з тексту. Так, не так давно британська компанія Google DeepMind представила

новий алгоритм для синтезу людського мовлення під назвою WaveNet. В його основі лежить використання нейромереж, що дозволяє домогтися більш реалістичної імітації голосу. Робота алгоритму WaveNet полягає в поточній генерації профілю звукової хвилі для кожного часового відліку за допомогою спеціальної нейромережі. Хоча навчання мережі проводилося за допомогою записів голосу диктора, система не використовує але звукові фрагменти цієї бази, а генерує їх самостійно. При створенні програми розробники використовували нейромережу типу FCN (повністю згортова мережа), архітектура якої була натхненна нейромережами PixelRNN та PixelCNN. Кожен згортковий шар в цій мережі має свій множник розширення, завдяки якому її рецептивне поле, тобто частина інформації, яку обробляють нейрони, зростає експоненціально. Так, шар, котрий іде за вхідним, має множник розширення, що дорівнює одиниці. Наступний множник дорівнює вже двом, множник за ним - чотирьом, а множник вихідного шару дорівнює восьми (Рис 2) . По суті, це дозволяє програмі охоплювати відразу велику кількість часових кроків. У нейромережі також передбачено зворотний зв'язок, тому кожен наступний звук машинної мови генерується на основі безлічі попередніх. Як повідомляють розробники, WaveNet може пам'ятати до 2-3 попередніх фонем [30]. У 2017 році Google представив архітектуру нової нейромережі Takotron2 [21], яка базується на вокодері WaveNet [40], а роком раніше на основі WaveNet розробила першу професійну систему синтезу української мови [24].

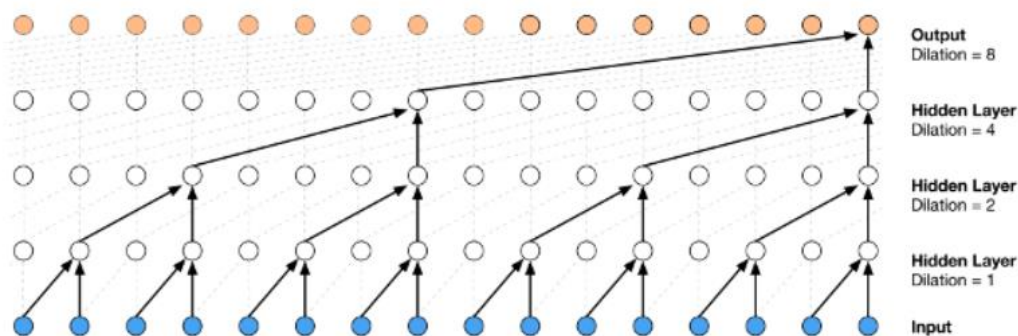


Рис 2. Aaron van den Oord et al. WaveNet

Озвучування довільного тексту вимагає вирішення низки завдань, пов'язаних із його попередньою лінгвістичною обробкою. Тому в системах автоматичного синтезу мови з'явився лінгвістичний етап обробки синтезованих повідомлень. Сучасні синтезатори мови типу "Текст-Мовлення" включають два блоки: лінгвістичної обробки (англ. Natural Language Processing, NLP), за допомогою якого для будь-якого довільного повідомлення або тексту будується його повна фонетична транскрипція, що включає як сегментну, так і просодичну характеристики; та акустичного синтезу або цифрової обробки сигналу (англ. Digital Signal Processing, DSP), що генерує мовний сигнал, тобто "озвучує" фонетичну транскрипцію.

О. Кривнова зазначає, що «Головним завданням блоку лінгвістичної обробки є створення транскрипційного уявлення (запису) тексту і його фрагментів, яке повинно містити як сегментну, так і суперсегментну фонетичну інформацію, необхідну для подальшої генерації акустичного сигналу» [11].

І.С. Кипяткова та А.А. Карпов аргументують перевагу саме автоматично генерованих транскрипцій для систем автоматичного розпізнавання та синтезу мовлення тим, що «транскрипції можуть бути створені вручну, але цей процес є трудоемким, а при розробці нової системи з новим набором слів транскрипцію необхідно створювати заново» [10].

Для систем розпізнавання мовлення транскрипції створюються на основі списку слів, а для систем синтезу мовлення – на основі вхідного тексту. При створенні транскрипції тексту необхідно брати до уваги фонетичні явища, що відбуваються на стику слів. Більш того, як зазначають І.С. Кипяткова та А.А. Карпов у своєму дослідженні розпізнавання та синтезу мовлення, «при створенні транскрипції для систем розпізнавання розмовного мовлення для кожної словоформи необхідно створювати альтернативні транскрипції, які враховували б різноманітні можливі варіанти вимови слів, що враховують усі можливі варіанти вимови одного й того ж слова в розмовному мовленні» [10].

Альтернативні транскрипції, на думку вчених, є необхідними при розробці систем розпізнавання та синтезу розмовного мовлення. У розмовному мовленні вимова слів може мати свої варіанти: різні диктори мають різні способи вимови того самого слова, більш того, вимова того самого диктора може змінитися залежно від швидкості мовлення та контексту [10]. Розмовне мовлення, серед інших, представлено фонетично інтонаційними особливостями [6]. У своїй статті Антоніна Чернявська та Віра Навроцька стверджують, що на сьогодні результати багатьох експериментально-фонетичних досліджень дають підставу зробити висновок про те, що набір власних фонетично-просодичних засобів, спрямованих на реалізацію комунікативних завдань, є притаманним кожному окремому функціональному стилю [34].

Розмовному мовленню притаманні такі явища, як асиміляція, у багатьох мовців - редукція деяких фонем до їх зникнення, оглушення приголосних у складі префіксів на межі префікса та кореня, оглушення приголосних звуків у кінці слів [20] тощо. Саме тому транскрипції слів, що вживаються в розмовному мовленні, у текстовому форматі часто не співпадають з транскрипціями, зробленими за нормативними літературними правилами української мови. Наприклад, слово «Здрастуйте!», що має базову транскрипцію [здра`стуйте], у сучасному розмовному мовленні є фонетично видозміненим й часто вимовляється як [здра`ст`е] або [дра`ст`і] [13]. Або ж слово шістдесят, що у розмовному мовленні зустрічається як нормативне [шіздес`а`т] та ненормативне [шийс`а`т]. Саме для врахування редукції та асиміляції є необхідним включення альтернативних транскрипцій до словника системи розпізнавання при автоматичному аналізі чи синтезі розмовного тексту.

Висновки до розділу 2

Отже, розробка автоматичної фонетичної транскрипції тексту лежить в основі створення якісних технологій синтезу тексту, дозволяючи економити час та ресурси та надаючи більшу варіативність. При створенні таких транскрипцій необхідним є врахування фонетичних та інтонаційних особливостей мовлення диктора та включення альтернативних варіантів вимови слів.

РОЗДІЛ 3. Створення автоматичного генератора фонетичної транскрипції для українськомовних текстів

У цьому розділі ми опишемо процес розробки генератора транскрипцій, принципи роботи алгоритму та правила транскрибування, на яких він базується. Також представимо процес тестування й виправлення помилок готової програми та окреслимо плани й перспективи покращення роботи транскриптора.

3.1. Розробка транскриптора для українськомовних текстів

Транскриптор створено в інтегрованому середовищі розробки PyCharm IDE (Integrated Development Environment) мовою програмування Python. Розробка програми проходила в декілька етапів. Перший етап включав у себе створення стандартної фонематичної транскрипції, що працювала б із символьними послідовностями в межах окремих слів. Це завдання було досягнуто шляхом задання списку правил для перетворення текстових символів. Формувалися списки звуків, поділених за групами голосні – приголосні, м'які – напівпом'якшені тощо, до яких далі звертався алгоритм, як і списки префіксів для вирішення проблеми відображення буквосполучень «дж» та «дз» як двох окремих звуків на стику морфем. Програма тестувалася на окремих словах та не враховувала фонетичні зміни на стику слів.

Наступним етапом розробки стало включення до алгоритму акомодацийних змін, що повело за собою ряд збоїв у роботі програми в результаті накладання правил одне на одне. На даному етапі розробки стало необхідним ієрархічне структурування коду. Проблему було вирішено розміщенням ряду правил у певній заданій послідовності. Спочатку відбуваються всі перетворення символів, пов'язані з позначенням букв та буквосполучень більшою чи меншою кількістю звуків у транскрипції. Наступними розміщено правила, результатами яких є перетворення звуків на інші: наприклад, унаслідок асиміляції. На цьому етапі роботи алгоритму

запускається наступний цикл програми, який працює з текстом із заміненними в попередніх описаних кроках символами. Тепер обробляється ряд правил, який включає в себе заміну символів на діакритичні знаки (м'який знак або апостроф на знаки м'якості та напівпом'якшення) та додає діакритичні знаки на позначення таких фонетичних змін як лабіалізація, палаталізація та і-подібна артикуляція.

Додавання кількох діакритичних знаків до одного звуку також відбувається в певній послідовності. На Рис.3 показано результат транскрибування слова «телебаченню», де позначення комбінаторних змін у вимові приголосного звуку записуються в такій послідовності: спочатку відбувається відображення м'якості, після цього – подовження, і аж потім позначається огублення.

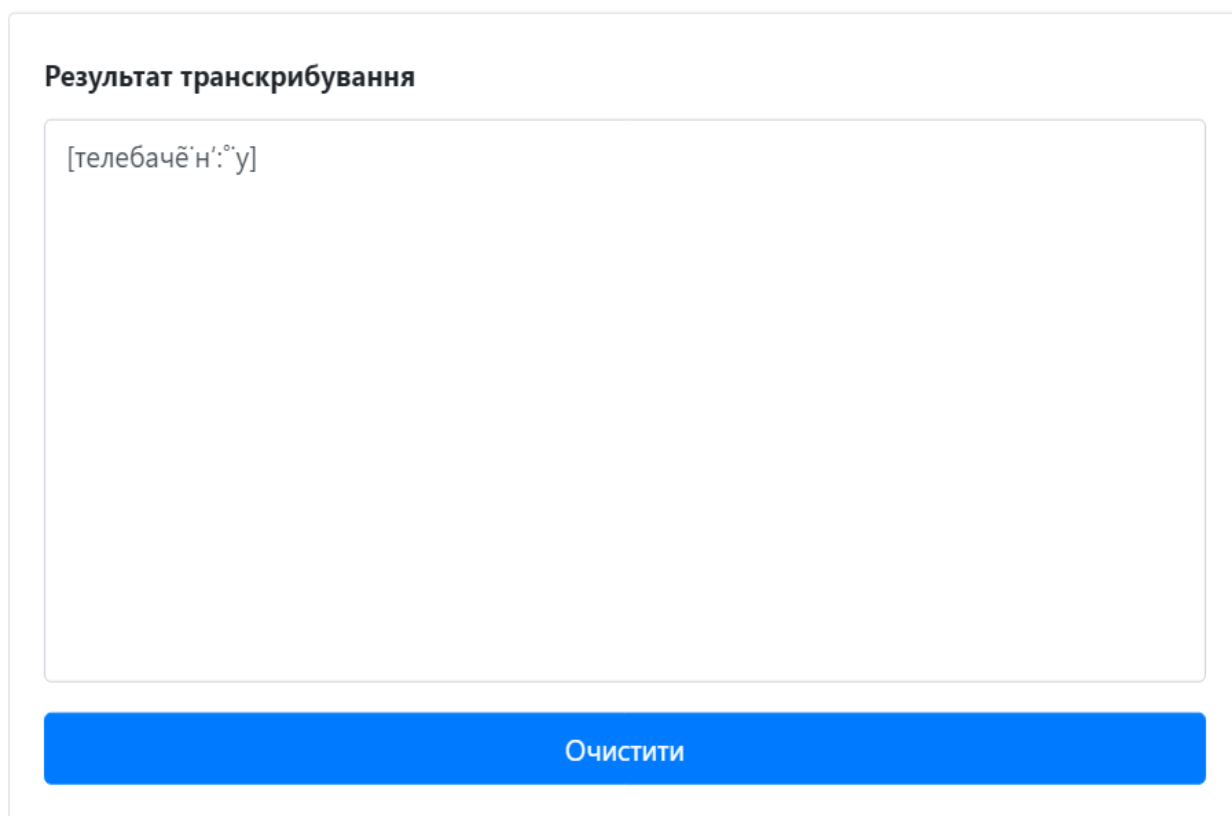


Рис. 3. Послідовність відображення комбінаторних змін у роботі транскриптора

Такий результат обумовлений ієрархією правил перетворення звуків унаслідок комбінаторних змін. У буквосполученні «нню» за заданими правилами відбувається такий ряд перетворень:

1. перш за все нотується м'якість приголосного [н'] перед «ю», що позначає два звуки;
2. далі перший «н» розпізнається як ідентичний до свого наступного символу, і послідовність ««однаковий символ» + «однаковий символ»» замінюється на ««символ» + «:»»;
3. символи у послідовності ««:» + «'»» міняються місцями;
4. позначення у транскрипції лабіалізації приголосної перед губним «у», з урахуванням попередньо доданих діакритичних знаків при аналізі символічних послідовностей.

Другий крок представленого вище алгоритму відрізняється від послідовності змін, яку прийнято враховувати при транскрипції слів (позначення м'якості [н] перед «ю», що позначає два звуки; асиміляція попереднього звука [н] до наступного за м'якістю; позначення подовження м'якого [н']); проте вказані кроки є рудиментарними коли мова йде про алгоритм роботи програми, тому було прийнято рішення скоротити кількість кроків до описаних вище для підвищення ефективності коду з одночасним отриманням однакового результату.

На третьому етапі розробки було вирішено проблему відображення фонетичних змін на стику слів. Для цього символ пробілу був включений до правил аналізу та перетворення символічних послідовностей, з попередньо доданою процедурою усунення повторюваних пробілів та інших небажаних пробільних символів, які могли б порушити правильність роботи алгоритму. Також на даному етапі певною мірою було реалізовано відображення фонетичних слів та позначення інтонаційних пауз: створено списки слів, що належать до службових частин мови, які зливаються з наступним словом під час вимови та розроблені правила заміни розділових знаків з використанням регулярних виразів. Результат представлення фонетичного слова у звуковому

записі та взаємодію декількох правил перетворень можна побачити на прикладі транскрибування введеного сполучення прийменника з іменником

The image shows two screenshots of a transcription interface. The top screenshot is titled "Текст для транскрибування" (Text for transcription) and contains the text "З жінкою." where the word "жінкою" is underlined with a red dashed line. Below the text is a green button labeled "OK". The bottom screenshot is titled "Результат транскрибування" (Transcription result) and shows the phonetic transcription "[ж:ʔнк°оʔ°у//]" in a box.

«з жінкою» (Рис.4), розділеного пробілом, де прийменник [з], ігноруючи пробіл, приєднується до наступного слова та асимілюється до звука [ж] за способом творення. Готова транскрипція була заключена в квадратні дужки.

Рис. 4. Позначення фонетичних слів у результаті роботи транскриптора.

Наразі готовий алгоритм роботи програми починається з попередньої обробки поданого на вхід тексту, яка включає в себе такі кроки:

1. Відбувається приведення регістру, тобто перетворення усіх буквених символів, до нижнього регістру (Рис.5).

2. Прибираються всі повторювані пробільні символи, лапки та дефіси.

3. Повторювані розділові знаки, такі як крапки, двокрапки, знаки оклику та знаки запитання («...», «!!!», «???»), зводяться до одного символу (Рис.5). Розділові знаки замінюються на скісну риску (бекслеш) таким чином: «,», «-», «—», «:», «;» - «/», «.», «!», «?» - «//» . Програма не плутає дефіс, використаний у якості тире, та дефіс, що використовується для орфографічного запису слова, оскільки бере до уваги оточуючі символи (Рис.5).

Дівчина-красуня)))) ,,,, хлопець-багатир!!!!!!??

Назад

Надіслати

Результат транскрибування

[д'іўчїнā крас'ўн'а// хл'опе'ц' багатир//]

Рис 5. Результат роботи транскриптора

5. За допомогою використання регулярних виразів шляхом заміни символів реалізовано явище спрощення у таких групах приголосних: «стс» - «сс», «стц» - «сц», «стч» - «сч», «нтст» - «нст», «стськ» - «ськ». Спрощення не позначається у таких словах-винятках: «кістлявий», «пестливий», «хвастливий», «хвастнути», «зап'ястний», «хворостняк». (Рис. 6)

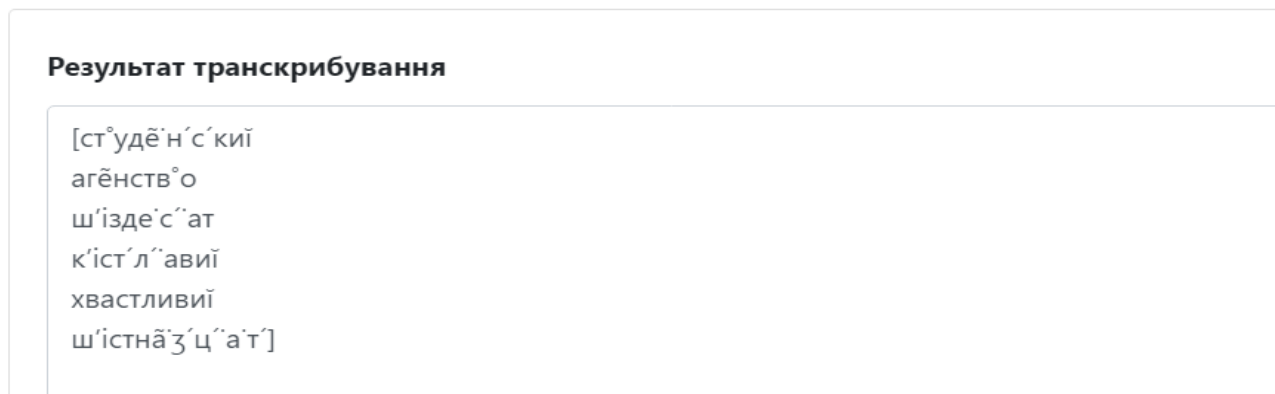


Рис. 6. Результат роботи транскриптора

Позначено оглушення приголосної «г» перед наступним глухим приголосним у результаті регресивної асиміляції у словах винятках: «кігті» - «кіхті», «нігті» - «ніхті», «легко» - «лехко» - «вогко» - «вохко», «дъогтю» - «дюхтю». (Рис. 7)

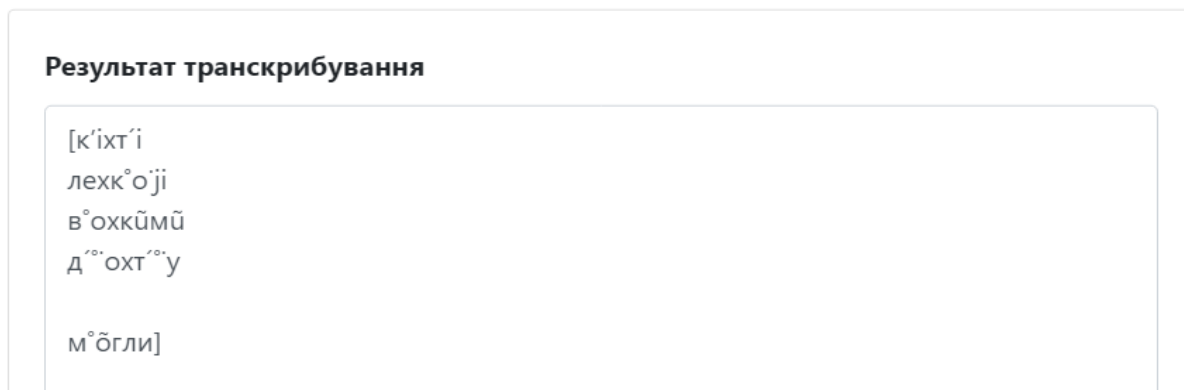


Рис. 7. Результат транскрибування

7. Програма розрахована на врахування фонетичних слів. Для цього були створені списки службових частин мови, до яких звертається алгоритм. Програма прибирає пробільні символи перед потрібним словом або ж після нього. Таким чином, формотворча частка «би», що утворює умовний спосіб дієслова, приєднується до попереднього слова: «сказав би» – «сказавби», а частка «хай», що утворює наказовий спосіб дієслова, – до наступного: «хай скаже» – «хайскаже». При цьому слова, що мають омонімічні відповідники серед інших частин мови (прийменник «коло» – іменник «коло», прийменник «навколо» - прислівник «навколо»), не були включені до списку. (Рис. 8)

Позначте наголоси (опціонально)

на майбутнє без жалю
міг би не може
хоча б

Назад

Надіслати

Результат транскрибування

[nāmāib˚ut˚n˚ē bej:ɑ˚l˚u
m˚ıgbı nēm˚ōje
x˚oçɑb]

Рис.8. Результат роботи транскриптора

Після здійснення попередньої обробки тексту програма виконує алгоритм дистрибутивного аналізу текстових символів та їх оточення. Серед символів тексту на цьому етапі залишаються літери, цифри та пробіл. Програма здійснює перетворення лише кириличних символів, що входять до українського алфавіту, тому звукове відображення слів, записаних символами латинського алфавіту (якщо такі трапляються у вхідному тексті), не подається у фонетичному записі на виході.

Алгоритм працює за принципом аналізу символічних послідовностей та подальшої їх заміни на відповідні фонетичні еквіваленти. Приклад роботи алгоритму представлено на Рис.9. Тут списком правил задано представлення регресивної асиміляції приголосних звуків за м'якістю: програма перевіряє приналежність символу до списку звуків, що можуть уподібнюватися за м'якістю до наступних м'яких та звуків, що можуть уподібнюватися за м'якістю до наступних напівпом'якшених. Далі аналізуються символічні послідовності довжиною в два символи, і якщо вони співпадають з патерном «м'який + «'»» чи «напівпом'якшений + «'»», відбувається відповідне додавання символу м'якості чи напівпом'якшення до патерну.

```
for symb in range(len(text) - 1):  
  
    if text[symb] in soft[1:] and (text[symb + 1] + text[symb + 2]) in softy:  
        text = text.replace(text[symb] + text[symb + 1] + text[symb + 2],  
                             text[symb] + "'" + text[symb + 1] + text[symb + 2])  
    if text[symb] in soft[5:] and (text[symb + 1] + text[symb + 2]) in semisoft[8]:  
        text = text.replace(text[symb] + text[symb + 1] + text[symb + 2],  
                             text[symb] + "'" + text[symb + 1] + text[symb + 2])  
    if text[symb] in soft[5:] and (text[symb + 1] + text[symb + 2]) == 'p' + "'":  
        text = text.replace(text[symb] + text[symb + 1] + text[symb + 2],  
                             text[symb] + "'" + text[symb + 1] + text[symb + 2])
```

Рис.9. Приклад алгоритму, заданого правилами позначення регресивної асиміляції за м'якістю

Результат виконання алгоритму представлений на Рис. 10. Як бачимо, звук [л] уподібнюється до наступного м'якого приголосного [л'], що відображено у транскрипції діакритичним знаком «'», [с] уподібнюється до напівпом'якшеного [м'], а [з] – до [р'].

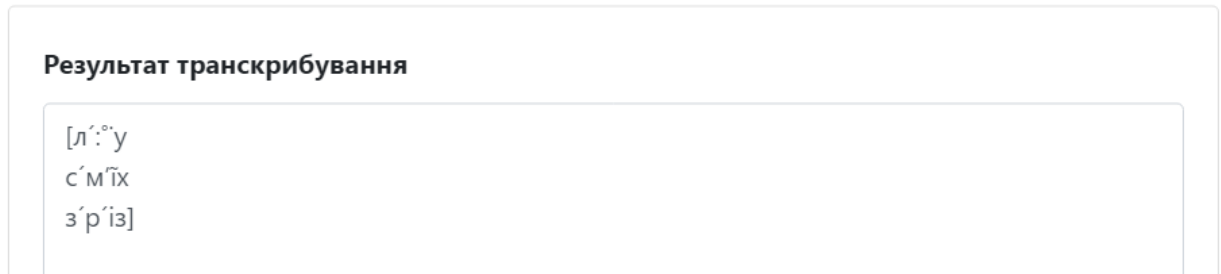


Рис.10. Результат роботи фонетичного транскриптора на прикладі відображення регресивної асиміляції за м'якістю

3.2. Правила транскрибування

На даному етапі розробки програма обробляє такі комбінаторні модифікації приголосних:

1. Акомодації:

- а) напівпом'якшення;
- б) лабіалізація.

2. Асиміляції:

- а) за дзвінкістю;
- б) за глухістю;
- в) за способом творення;
- г) за місцем творення;
- г) за місцем і способом творення;
- д) за м'якістю.

3. Подовження.

4. Спрощення.

У транскрипції також відображено наступні акомодативні зміни голосних звуків:

- а) назалізація (прогресивна та регресивна);
- б) і-подібна артикуляція (прогресивна та регресивна).

Перетворення символів відбувається за такими правилами алгоритму:

1. Літери «ї» та «щ» у транскрипції завжди передаються двома звуками: [j^і] та [шч] відповідно (Рис.11).

2. Голосні «я», «ю», «є» передаються двома звуками як [j^а], [j^у], [j^є] у таких позиціях:

- а) на початку тексту;
- б) після голосного;
- в) після апострофа;
- г) після м'якого знаку. (Рис.11)

3. Перед приголосними голосні «я», «ю», «є» передаються одним звуком як [а], [у], [є].

4. Звук [й] записується як [j] при транскрибуванні «ї» та «я», «ю», «є», що позначають два звуки (Рис.11).

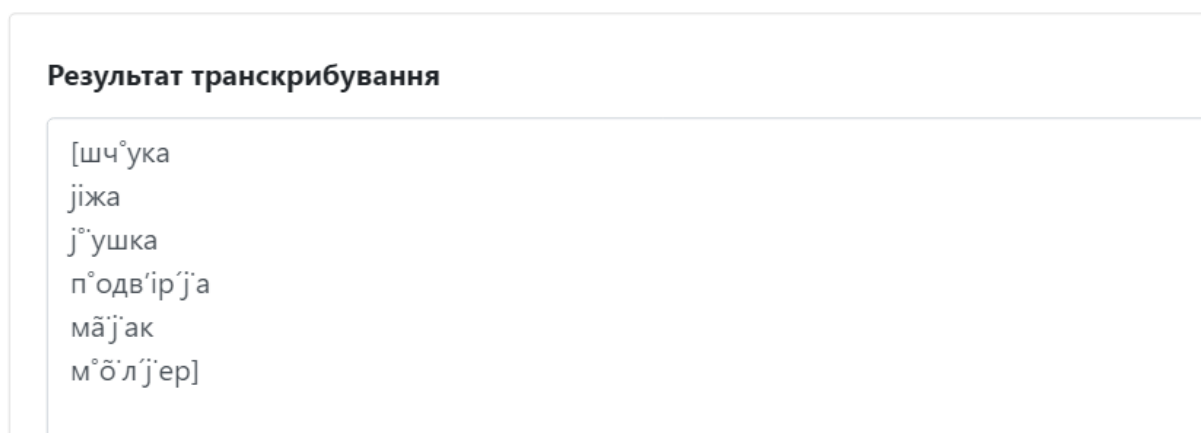


Рис. 11. Результат роботи транскриптора

5. Звук [в] вокалізується, перетворюючись на нескладовий [ў] у таких позиціях:

- а) на початку тексту перед приголосним;
- б) усередині слова після голосного перед приголосним;

в) у кінці тексту після голосного (Рис.12).

6. Звук [й] вокалізується, перетворюючись на нескладовий [i̯] у таких позиціях:

- а) на початку тексту перед приголосним;
- б) усередині слова після голосного перед приголосним;
- в) у кінці тексту після голосного (Рис.12).

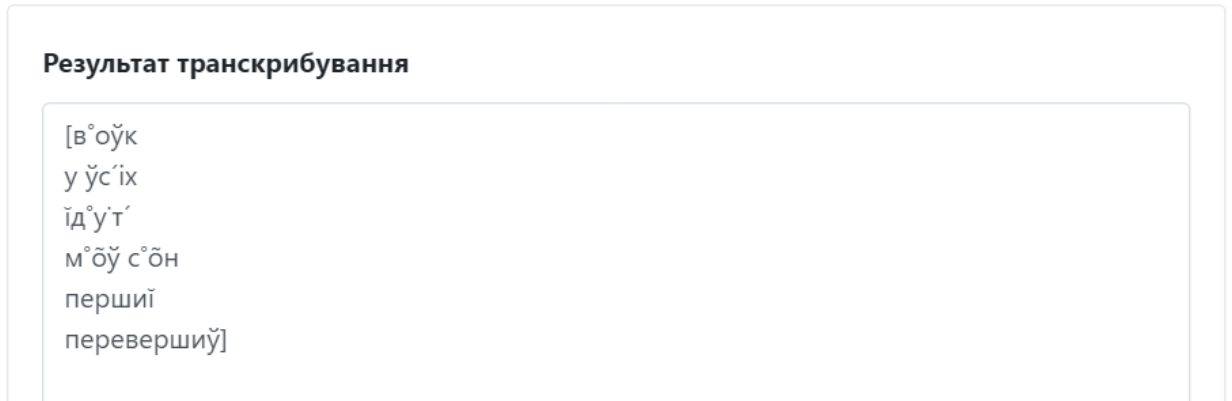


Рис. 12. Вокалізація у результаті роботи транскриптора

7. Буквосполучення «дж» та «дз» відображаються відповідними символами [ДЖ] та [З] входячи до складу однієї морфеми та як [дж] та [дз] знаходячись на межі префікса й кореня. (Рис.13)

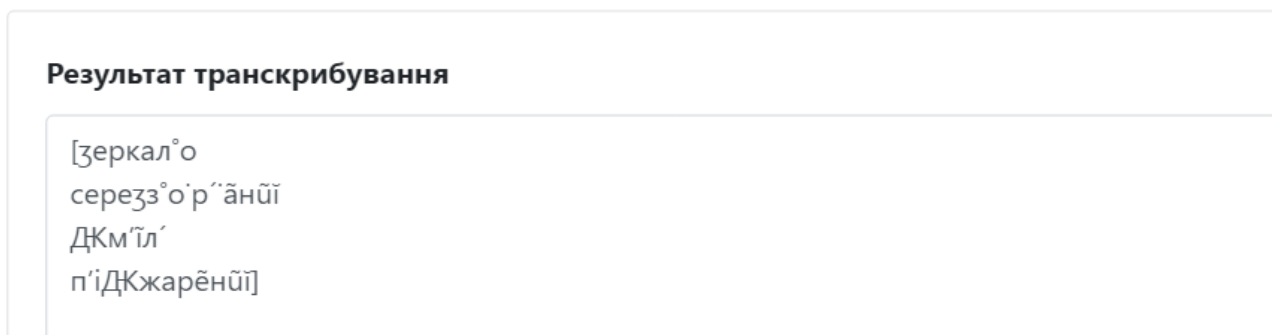


Рис.13. Транскрибування “дж” та “дз”

8. М'які звуки відображаються як [д'], [т'], [з'], [с'], [ц'], [л'], [н'], [з'], [р'] перед такими символами:

- а) «я», «ю», «є», що позначають один звук;
- б) м'який знак «ь» або «і» (Рис.14).

9. Шиплячі [ж], [ДЖ], [ч], [ш], губні [б], [п], [в], [м], [ф], глотковий [г] та задньоязикові [к], [х], [г] пом'якшуються та відображаються як [ж'], [ДЖ'], [ч'], [ш'], [б'], [п'], [в'], [м'], [ф'], [г'], [к'], [х'], [г'] перед «і» та «я», «ю», «є», що позначають один звук (Рис.14).

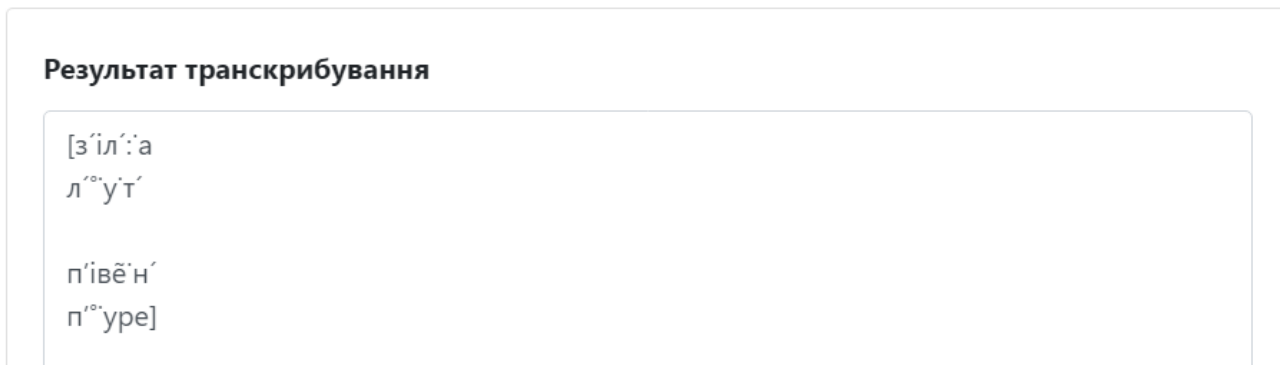


Рис.14. Результат роботи транскриптора

10. Глухі приголосні [п], [т], [ц], [с], [к], [ч] регресивно одзвінчуються перед дзвінками [б], [д], [з], [з], [г], [г], [ДЖ], [ж] та відповідно перетворюються на [б], [д], [з], [з], [г], [ДЖ]. (Рис.15)

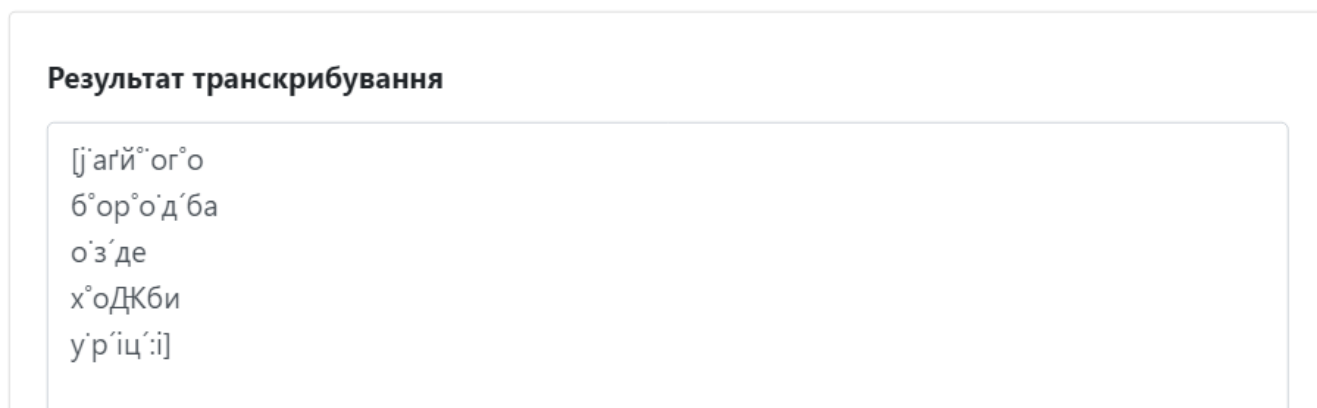


Рис.15. Результат роботи транскриптора

11. Приголосні [д], [т], [ж], [ш], [ч] уподібнюються до наступних свистячих та перетворюються відповідно на [з], [ц], [з], [с], [ц] (Рис).

12. Приголосні [д], [т], [с], [з] уподібнюються до наступних шиплячих та перетворюються відповідно на [ДЖ], [ч] [ш], [ж] (Рис.16).

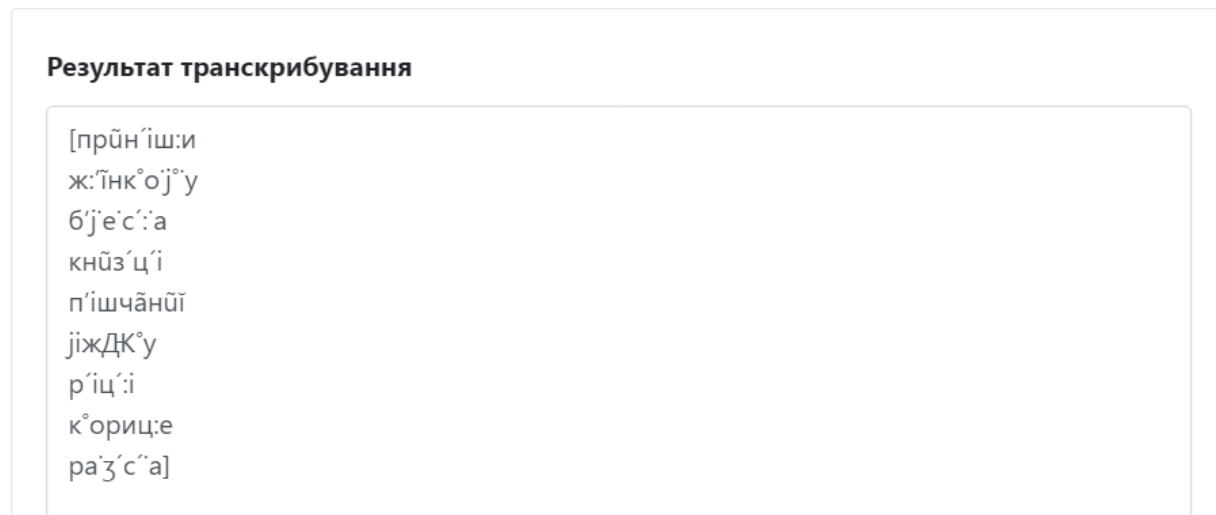


Рис.16. Результат роботи транскриптора

13. Буквосполучення «ться» перетворюється на «тьця» (та відображається як [ц':`а] у готовій транскрипції) (Рис.17).

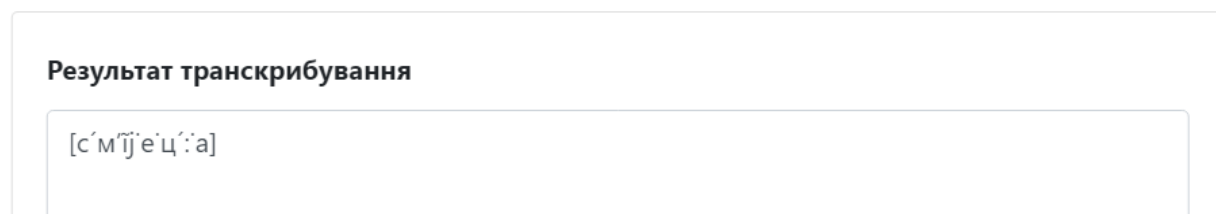


Рис. 17. Результат роботи транскриптор

14. Приголосні [д], [т], [з], [с], [ц], [л], [н], [з] уподібнюються за м'якістю до наступних м'яких приголосних [д'], [т'], [з'], [с'], [ц'], [л'], [н'], [з'] та перетворюються на [д'], [т'], [з'], [с'], [ц'], [л'], [н'], [з']. (Рис.18)

15. Приголосні [з], [з], [с], [ц] уподібнюються за м'якістю до наступних напівпом'якшених шиплячих [ж'], [ДЖ'], [ч'], [ш'], губних [б'], [п'], [в'], [м'], [ф'], глоткового [г'] та задньоязикових [к'], [х'], [г'] та м'якої [р'] перетворюючись на [з'], [з'], [с'], [ц']. (Рис.18)

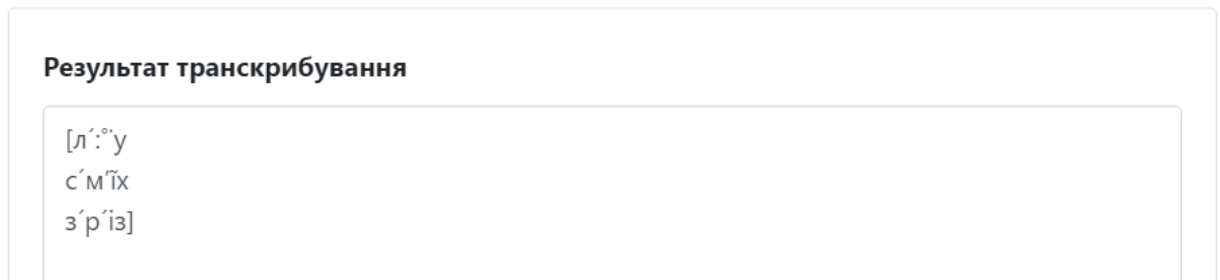


Рис. 18. Результат роботи транскриптора

16. Перед голосними [o] та [y] приголосні регресивно лабіалізуються, що в транскрипції позначається додаванням символу [°] до приголосного звуку: [к°o]. (Рис.19)

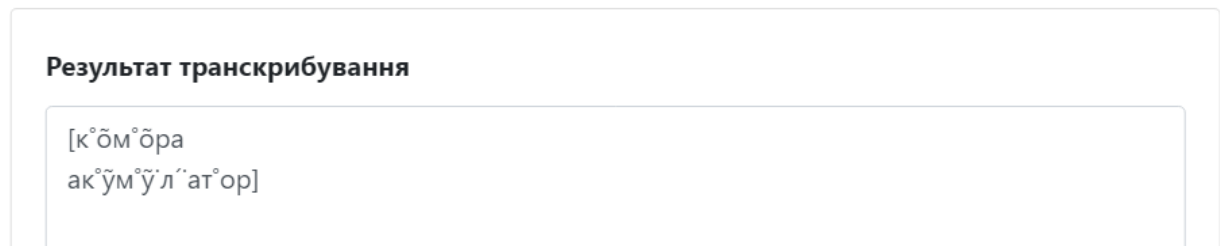


Рис. 19. Результат роботи транскриптора

17. Поруч із приголосними [н], [н'] та [м] голосні звуки набувають прогресивної та регресивної назалізації, що в транскрипції позначається додаванням хвильки над голосним звуком: [мãm'ї]. (Рис.20)

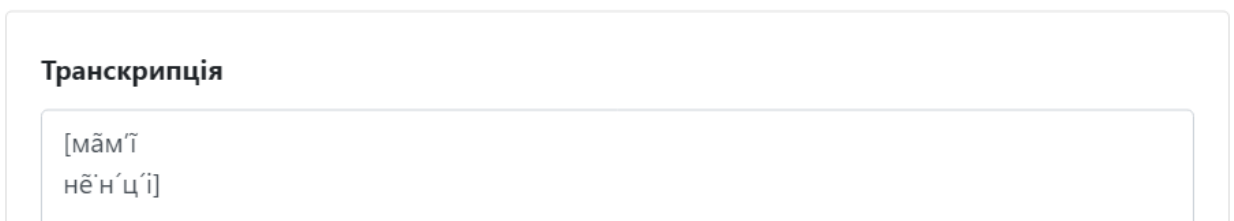


Рис. 20. Результат роботи транскриптора

18. Поруч із м'якими приголосними [д'], [т'], [з'], [с'], [ц'], [л'], [н'], [з'], [р'], [й], [j] голосні звуки [а], [е], [о], [у] набувають прогресивної та регресивної і-подібної артикуляції, що в транскрипції позначається додаванням символу ['] до голосного звуку: [л'а'л'ка] (Рис.21).

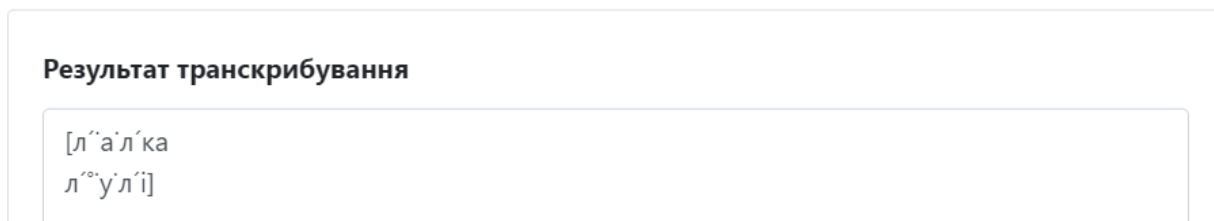


Рис. 21. Результат роботи транскриптора

Результати транскрибування уривку із роману Пантелеймона Куліша «Чорна рада» представлені у Додатку 1, Додатку 2 та Додатку 3.

3.3. Тестування розробленої програми транскрибування текстів

Програма тестувалася на масивах текстів різних функціональних стилів. У ході тестування нам зустрічалися помилки у вихідній транскрипції, пов'язані як з неповністю правил обробки тексту, так і з наявністю у вхідному тексті нестандартних символів Юнікод.

Наприклад, в уривку, продемонстрованому на Рис. 22, знак апострофа, використаний у тексті, не відповідає знаку апострофа, занесеного до алгоритму; це призвело до невиконання правила відображення голосної «ю» за допомогою двох звуків та заміну апострофа на пустий символ.

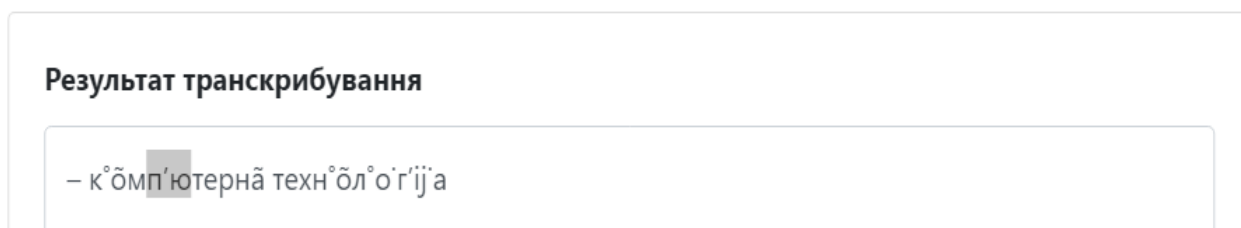


Рис. 22. Результат роботи транскриптора

Проблему було вирішено заміною алгоритмом програми проблемного символу на стандартний апостроф за допомогою регулярних виразів. (Рис.23)

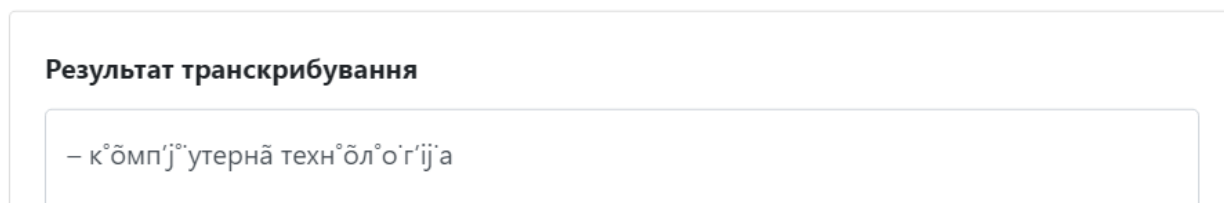


Рис.23. Результат роботи транскриптора

Ще одна помилка, пов'язана із небуквеними символами, була помічена у сполученні закриваючої дужки з комою («),»). (Рис.24)

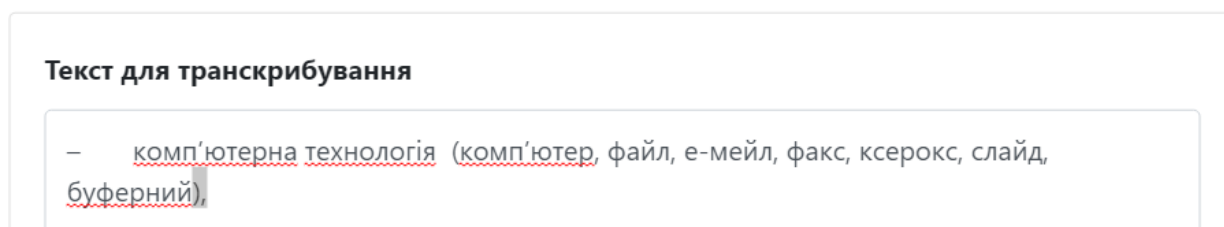


Рис.24. Результат роботи транскриптора

На Рис. 25 показано, як не спрацювало правило вокалізації «й» після паузи перед приголосним звуком, тому що правилами не було задано врахування символу переносу на наступний рядок.

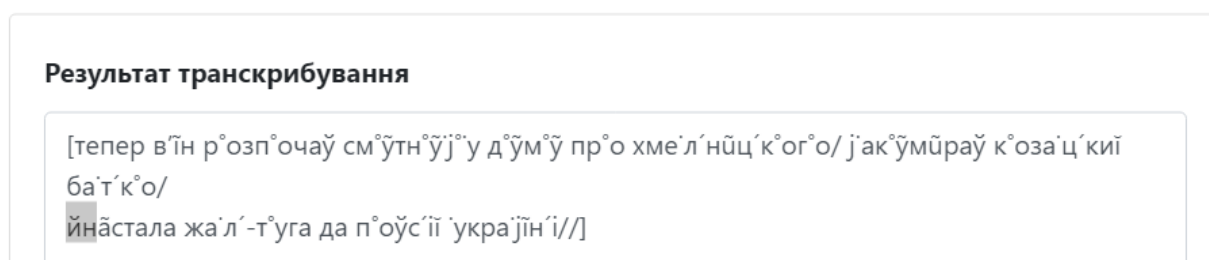


Рис. 25. Результат роботи транскриптора

Помилку було виправлено редагуванням правил алгоритму. (Рис.26)

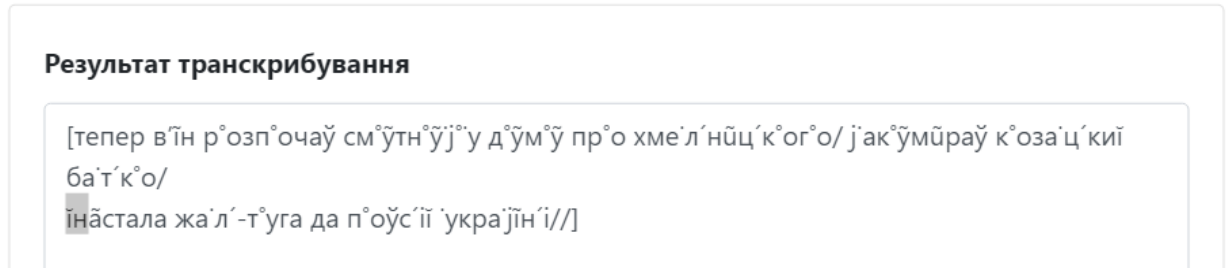


Рис.26. Результат роботи транскриптора

Ще одним багом, із яким ми зіткнулися, було неврахування можливості м'якості попереднього приголосного правилами усунення дефісу між двома частинами складного слова. (Рис. 27)

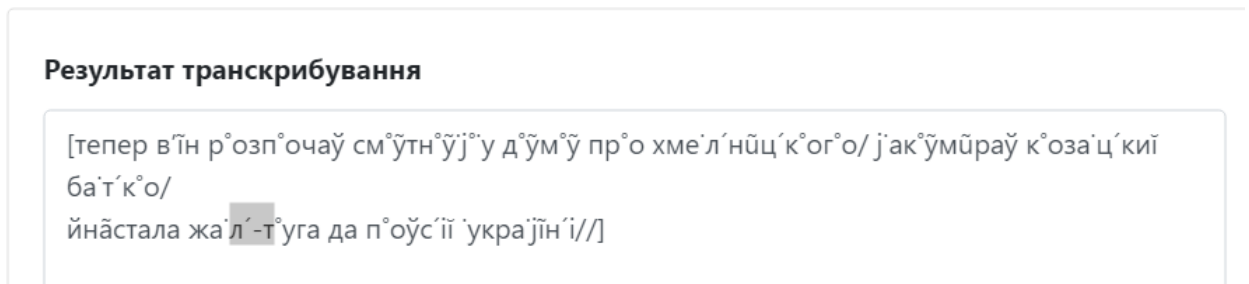


Рис.27. Результат роботи транскриптора

Помилку було виправлено редагуванням правил алгоритму. (Рис. 28)

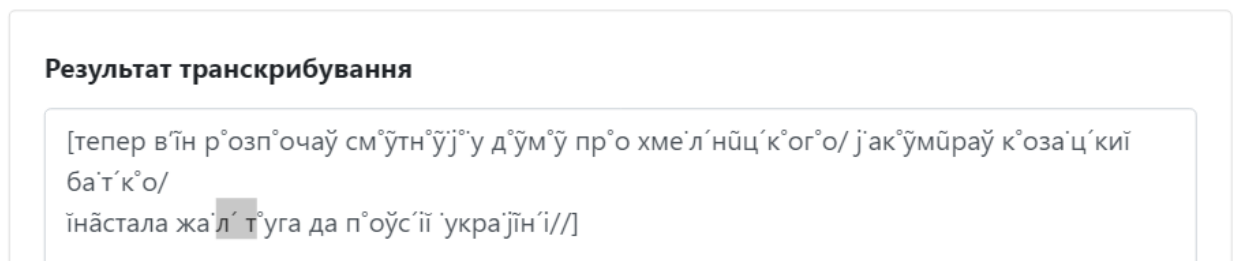


Рис.28. Результат роботи транскриптора

Після знаходження помилки, вказаної на Рис.29, до правил було додано відображення літер на позначення двох звуків двома символами після попереднього м'якого приголосного, м'якість котрого позначена на письмі м'яким знаком у словах іншомовного походження (Рис. 30).

Результат транскрибування

[н'юфаўндлєнд]

Рис. 29. Результат роботи транскриптора

Результат транскрибування

[н'ј'уфаўндлєнд]

Рис. 30. Результат роботи транскриптора

При спробі позначити наголошені голосні великою літерою ми стикнулися з проблемою не розпізнавання правилами лабіалізації великих літер.верхнього регістру. (Рис.31)

Рис.31. Результат роботи транскриптора

Результат транскрибування

[с°олОнка]

Проблему було вирішено оновленням правил огублення несприятливими до регістру (Рис.32).

Результат транскрибування

[с°ол°Онка]

Рис.32. Результат роботи транскриптора

Як видно, та сама помилка розповсюджується й на назалізацію.

У транскрипції також зустрічаються випадкові хиби, пов'язані не з помилкою у правилах, а радше зі збоєм у роботі алгоритму внаслідок взаємодії великої кількості операторів.

Так, на Рис. 33 показаний пропущений символ прогресивної назалізації /і/ під впливом попередньої /н/ близько кінця поданого на аналіз тексту. Схожі помилки можна спостерігати і с іншими змінами звуків у подібній позиції.

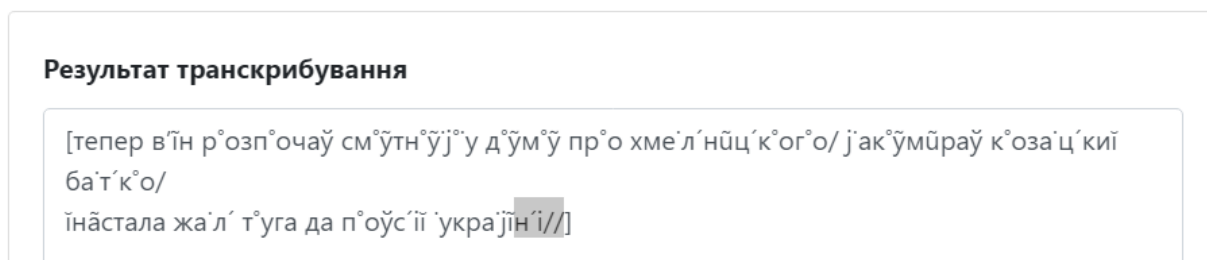


Рис.33. Результат роботи транскриптора

Так, на Рис.34 показаний пропущений символ і-подібної артикуляції перед напівпом'якшеним приголосним у об'ємному уривку тексту.

У той же час у меншому масиві тексту на тому ж місці алгоритм працює правильно. (Рис.35)

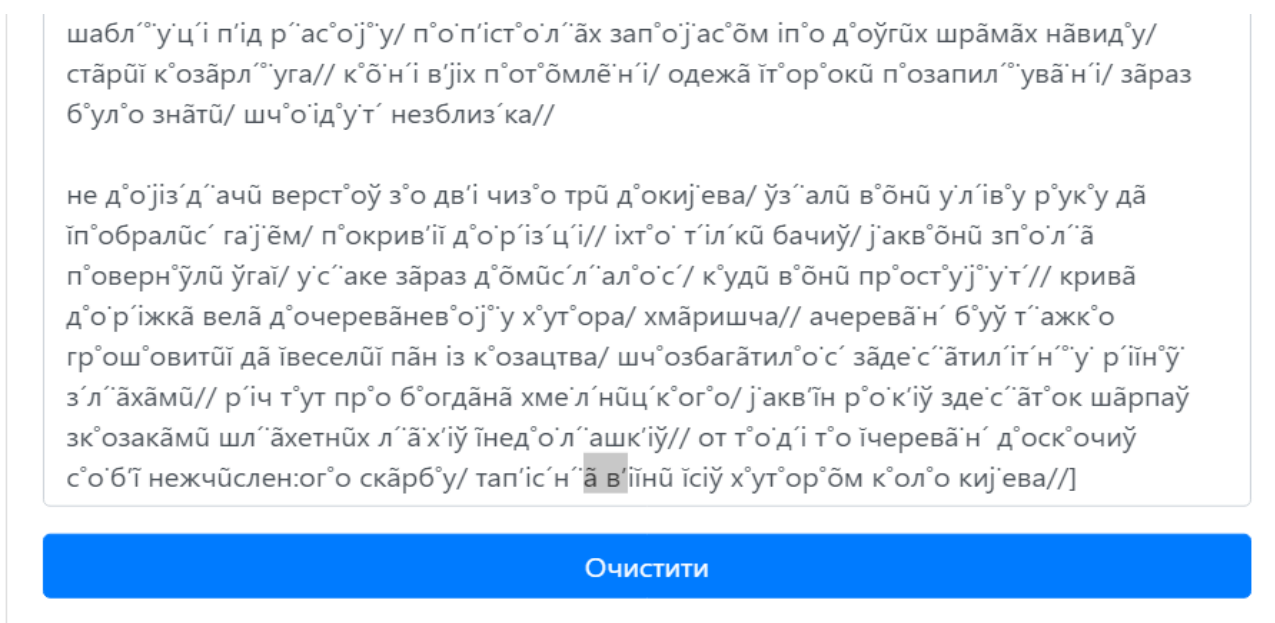


Рис.34. Результат роботи транскриптора

Результат транскрибування

[ачеревā́н' б'у́ т'ажк'о гр'ош'овити́ дā́ ї весели́ пā́н із к'озацтва/ шч'озбагатил'о'с' заде'с'атил'іт'н'о'у' р'і́йн'у́ з л'ахā́мї// р'іч т'ут пр'о б'огдā́нā́ хме'л'нїц'к'ог'о/ ј'акв'їн р'о'к'ї́у з де'с'ат'ок шарпа́у з к'озакā́мї шл'ахетнїх л'а'х'ї́у їнед'ол'ашк'ї́у// от т'од'і т'о ї черевā́н' д'оск'очи́у с'об'ї нежчислєн:ог'о скарб'у/ та'п'іс'н'ā́ в'ї́їнї́ ї с'ї́у х'ут'ор'ом к'ол'о кијева//]

Рис. 35. Результат роботи транскриптора

Схоже, що програма починає давати збої по досягненню певної кількості символів.

Існує також ряд неузгоджених правил стосовно позначень фонетичних слів, пов'язаних із неповнотою списків слів службових частин мови, що приєднувалися б до повнозначних, з'єднання слів службових частин мови в одне фонетичне слово та складністю урахування всіх символічних послідовностей та розділових знаків для правильної реалізації даної функції.

3.4. Плани та перспективи покращення автоматичного транскриптора

Перед нами стоїть низка завдань для удосконалення алгоритму:

1. Позначення часткового оглушення дзвінких приголосних на стику морфем між префіксом та коренем та границі прийменника з іменником.

2. Усунення асиміляції приголосних за м'якістю на стику морфем між префіксом та коренем.

3. Позначення наголосу в словах, що уможливить відображення явищ наближення голосних [e] до [и], [и] до [e] в ненаголошеній позиції та [o] до [y] перед складом з наголошеними [y] до [i].

Дані проблеми можуть бути вирішені шляхом підключення баз даних з інформацією про наголос слів та зняттям омонімії префіксів та коренів «роз», «без». На даний момент ведеться розробка можливості для

користувача проставляти наголос вручну. Такий спосіб не є продуктивним при роботі з великими масивами текстів, але є оптимальним при транскрибуванні одиничних слів чи невеликих речень.

Як було продемонстровано вище, наш транскриптор час від часу допускає випадкові помилки механічного характеру. Це пов'язано зі складністю структури коду та взаємодією великої кількості умовних операторів, що робить програму негнучкою та схильною до хиб. Наразі нами ведуться роботи, спрямовані на оптимізацію алгоритму, вирішення зазначених вище проблем та створення веб-сервісу для взаємодії з користувачем.

Висновки до розділу 3

Отже, нами розроблено програмний модуль з автоматичним генератором транскрипції, що дає можливість зробити представлення тексту українською мовою у фонетичному та фонематичному виглядах. Транскриптор працює з такими комбінаторними змінами звуків мовлення як асиміляція (за дзвінкістю, глухістю, м'якістю, способом та місцем творення), напівпом'якшення, лабіалізація, подовження та спрощення, а також такими акомодативними змінами голосних звуків як назалізація та і-подібна артикуляція. Крім того, передбачено врахування фонетичних слів.

ВИСНОВКИ

Хід і результати дослідження дали змогу зробити такі висновки:

1. На підставі комплексного аналізу наукової літератури обґрунтовано погляди різних дослідників на визначення фонем та її відтінків; досліджено розвиток транскрипції та способів відображення звуків мов різних народів світу; проаналізовано погляди вітчизняних дослідників на фонетичну, фонематичну, словофонематичну, морфофонематичну та практичну види транскрипцій. Обґрунтовано доцільність використання фонетичної та фонематичної транскрипцій у сучасній лінгвістиці.

2. Розглянуто особливості представлення української фонетичної системи в Міжнародному фонетичному алфавіті на основі аналізу праць мовознавців, виділено погляди сучасних науковців на сьогоденні проблеми транскрипції, зокрема питання усталеного звукового представлення слів, власних назв та термінів іншомовного походження.

3. Здійснено діахронічний аналіз розвитку технологій розпізнавання та синтезу мовлення, проаналізовано принцип роботи, недоліки й переваги різних синтезаторів TTS (Text-to-Speech), зокрема сучасних систем, оснований на методі конкатенативного синтезу та синтезаторів на основі нейромереж. Проаналізовано переваги використання навчання нейромереж для синтезу штучного мовлення; з'ясовано місце фонетичної транскрипції тексту у структурі роботи синтезаторів та обґрунтовано необхідність автоматизації процесу транскрибування в лінгвістичних блоках TTS.

5. Розроблено програмний модуль з автоматичним генератором транскрипції, що дає можливість зробити представлення тексту українською мовою у фонетичному та фонематичному виглядах. Транскриптор працює з такими комбінаторними змінами звуків мовлення як асиміляція (за дзвінкістю, глухістю, м'якістю, способом та місцем творення), напівпом'якшення, лабіалізація, подовження та спрощення, а також такими акомодативними змінами голосних звуків як назалізація та і-подібна артикуляція. Крім того, передбачено врахування фонетичних слів.

Подальша розробка програми передбачає подальший пошук помилок та оптимізацію коду; поповнення списків лексем, що беруть участь у творенні фонетичних слів; розробка альтернативних режимів транскриптора, які нотували б можливе оглушення дзвінких приголосних на стику морфем та/або укінці слів; а також створення бази даних з позначеними рухомими наголосами слів для автоматизації процесу відображення наголошення голосних. Крім того, планується розробка веб-сервісу для взаємодії з програмою.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Аванесов Р. І. Фонетика современного русского литературного языка / Р. І. Аванесов.. – 241 с.
2. Бакум П. З. Історія української фонології в підготовці майбутніх учителів-словесників / П. З. Бакум.
3. Бас-Кононенко О. Сучасна українська літературна мова. Лексикологія. Фонетика / О. Бас-Кононенко, А. Мойсієнко, Н. Плющ., 2010. – 270 с.
4. Вакуленко М.О. Наукові засади відтворювання запозичених та іншомовних слів: інваріантна транскрипція і транслітерація / Вакуленко М. О. // К.: Вісник книжкової палати. – 1999. – № 10. — 40 с.
5. Вакуленко М.О. Про переклад та запозичення чужомовних слів / Вакуленко М.О. // Вісн. НАН України. – 1995. – № 11-12. – С. 78 – 81.
6. Дудик П. С. Розмовно-побутовий стиль мови і мовлення [Електронний ресурс] / П. С. Дудик – Режим доступу до ресурсу: <http://litmisto.org.ua/?p=5456>.
7. Жовтобрюх, М. А. Курс сучасної української літературної мови / М. А. Жовтобрюх, Б. М. Кулик. - Київ : Рад. шк., 1965.
8. Кабберли П. Транслитерация и транскрипция: общие и специфические проблемы / Кабберли П. // Русский язык за рубежом // Институт русского языка имени А.С. Пушкина. – 1994. – № 2.– 135 с.
9. Калиев А. СИНТЕЗ РЕЧИ: ПРОШЛОЕ И НАСТОЯЩЕЕ / А. Калиев, С. Рыбин.
10. Кипяткова И. С. Модуль фонематического транскрибирования для системы распознавания разговорной русской речи / И. С. Кипяткова, А. А. Карпов. – Санкт-Петербург, 2008.
11. Кодзасов С. В. Общая фонетика / С. В. Кодзасов, О. Ф. Кривнова., 2001.
12. Кратко М. Проблеми транскрипції в укладанні англійськоукраїнських словників з математики, програмування та

обчислювальної техніки. [Електронний ресурс] / М. Кратко, Н. Кікець. – 2000. – Режим доступу до ресурсу: http://ena.lp.edu.ua:8080/bitstream/ntb/54625/2/2000n402_Kratko_M-Problemy_transkryptsii_v_99-100.pdf.

13. Лексикологія української мови як учення про лексичний склад [Електронний ресурс] – Режим доступу до ресурсу: https://subject.com.ua/lesson/mova/10klas_1/53.html.

14. Лобанов Б. М. Компьютерный синтез и клонирование речи / Б. М. Лобанов, Л. І. Цирульник. – Минск: "Белорусская наука", 2008.

15. Лобанов Б. М. Моделирование мультисловных и межсловных фонетико-акустических явлений полного и разговорного стилей речи в системе синтеза речи по тексту "Мультифон" / Б. М. Лобанов, Л. І. Цирульник..

16. Лобанов Б. М. Синтез речи в Минске (ретроспективный обзор и современное состояние) [Електронний ресурс] / Б. М. Лобанов – Режим доступу до ресурсу: https://ssrlab.by/wp-content/uploads/2014/01/2004_lobanov_sintez-rechi-v-minske.pdf.

17. . Маслов Ю.С. Введение в языкознание / Маслов Ю.С. – М., 1987. – 272 с.

18. Микульчик Р. Сучасні українські аббревіатури: перспективи і стан дослідження [Електронний ресурс] / Р. Микульчик. – 2005. – Режим доступу до ресурсу: <http://science.lpnu.ua/terminology/all-volumes-and-issues/visnik-no-538-2005/suchasni-ukrayinski-abreviaturi-perspektivi-i>

19. Мінковська І. Транскрибування і транслітерування: можливості та проблеми застосування в Україні [Електронний ресурс] / І. Мінковська – Режим доступу до ресурсу: <https://drive.google.com/file/d/1lBAwnU4GoVnqVniG1FqZLn5XPcuBXCvs/view>.

20. Оглушення приголосних звуків правило. Особливості читання приголосних [Електронний ресурс] – Режим доступу до ресурсу:

<https://kerchtt.ru/uk/oglushenie-soglasnyh-zvukov-pravilo-osobennosti-chteniya-soglasnyh-oglushenie/>

21. Преобразование текста в речь с помощью нейросети [Електронний ресурс] – Режим доступу до ресурсу: <https://www.reg.ru/blog/text-to-speech/>.

22. Реформатский А.А. Введение в языкознание. / Реформатский А.А. – М.: 1996. – 542 с.

23. Синтезатор мовлення «фонемафон – 2000» [Електронний ресурс] – Режим доступу до ресурсу: http://uiip.bas-net.by/structure/l_rsr/pr4.php.

24. Синтез мовлення [Електронний ресурс] – Режим доступу до ресурсу: https://uk.wikipedia.org/wiki/Синтез_мовлення.

25. Синявський О. Спроба звукової характеристики літературної української мови / О. Синявський // Наукові записки Харківської науково-дослідної кафедри мовознавства. –

26. Стеріополо о. Українська фонетична система у парадигмі Міжнародної Фонетичної Асоціації (МФА) / О. Стеріополо

27. Сучасна українська літературна мова. Лексикологія, Фонетика. – Київ: Знання, 2010

28. Тоцька Н. І. Сучасна українська літературна мова / Н. І. Тоцька. – Київ: Головне видавництво видавничого об'єднання «Вища школа», 1981.

29. Тоцька Н.І. Фонетика і фонологія // Сучасна українська літературна мова: Підручник / А.П. Грищенко, Л.І. Мацько, М.Я. Плющ та ін.; За ред. А.П. Грищенко. – 3-тє вид., допов. – К.: Вища шк., 2002. – С.22.

30. Уласович К. Нейросеть научилась реалистично имитировать речь человека [Електронний ресурс] / Кристина Уласович – Режим доступу до ресурсу: <https://nplus1.ru/news/2016/09/14/google-ai-voice>.

31. Фонетична і фонематична транскрипція [Електронний ресурс] – Режим доступу до ресурсу: <https://studopedia.info/7->

90168.html#:~:text=a)%20фонетична%20транскрипція%20такий%20спосіб,що
%20неістотні%20у%20фонологічному%20плані.

32. Фонематична та фонетична транскрипції [Електронний ресурс] – Режим доступу до ресурсу: https://studopedia.ru/2_91100_fonematically-ta-fonetichna-transkriptsii.html.

33. Фонетичний алфавіт. Особливості фонематичної транскрипції [Електронний ресурс] – Режим доступу до ресурсу: https://studopedia.com.ua/1_21206_fonetichniy-alfavit-osoblivosti-fonematically-transkriptsii.html.

34. Чернявська А. Фоностилестична складова як засіб вираження змісту та відтінків висловлювання/ А. Чернявська, В. Навроцька., 2012.

35. Шишминцев С. Newfon SAPE [Електронний ресурс] / Сергей Шишминцев. – 2018. – Режим доступу до ресурсу: <https://nvda.ru/uploads/newfon/document.html>.

36. Bootstrap (фреймворк) [Електронний ресурс] – Режим доступу до ресурсу: [https://ru.wikipedia.org/wiki/Bootstrap_\(фреймворк\)](https://ru.wikipedia.org/wiki/Bootstrap_(фреймворк)).

37. Cucchiarini C. Automatic Phonetic Transcription:An overvie / C. Cucchiarini, H. Strik..

38. JavaScript [Електронний ресурс] – Режим доступу до ресурсу: <https://ru.wikipedia.org/wiki/JavaScript>.

39. Jurafsky D. Automatic Speech Recognition and Text-to-Speech / D. Jurafsky, J. Martin // Speech and Language Processing (3rd ed. draft) / D. Jurafsky, J. Martin.

40. Rabiner L. Digital Speech Processing [Електронний ресурс] / L. Rabiner, Biing Hwang Juang. – 2003. – Режим доступу до ресурсу: <https://www.sciencedirect.com/topics/computer-science/speech-synthesis>

ДОДАТКИ

Додаток А

Вхідний текст

Текст для транскрибування

По весні 1663 року двоє подорожніх, верхи на добрих конях, ізближались до Києва з Білогородського шляху. Один був молодий собі козак, збрійний, як до війни; другий по одежі і по сивій бороді, сказать би, піп, а по шаблюці під рясою, по пістолях за поясом і по довгих шрамах на виду — старий "козарлюга". Коні в їх потомлені, одежа й тороки позапилювані: зараз було знати, що ідуть не зблизька.

Не доїздячи верстов зо дві чи зо три до Києва, взяли вони у ліву руку да й побрались гаєм, по кривій доріжці. І хто тільки бачив, як вони з поля повернули в гай, усяке зараз домислялось, куди вони простують. Крива доріжка вела до Череваневою хутора, Хмарища. А Черевань був тяжко грошовитий да й веселий пан.

OK

Фонематична транскрипція

Результат транскрибування

[по весні 1663 року двоє подорожніх/ верхи надобрих конях/ ізближализ докијева з
білогорозського шляху// один був молодий собі козак/ зброїний/ жагдов'їни/ другий
поодежі і посив'її бороді/ сказад'би/ піп/ апо шабл'уці під'асоју/ поп'їстол'ах
запојасом і подоўгих шрамах навиду/ старий козарл'уга// коні їїх потомлен'ї/ одежа
їтороки позапил'уван'ї/ зараз було знати/ шчо ідут' незблиз'ка//

недоїз'дачи верстоў зодві чизотри докијева/ вз'али вони ул'їву руку даїпобрализ'
гајем/ покрив'її дор'їз'ці// іхто т'їл'ки бачиў/ јаквони зпол'а повернули ўгаї/ ус'аке
зараз домис'л'алос'/ куди вони простујут// крива дор'їжка вела дочереваневого
хутора/ хмаришча// ачереван' був т'ажко грошовитий даївеселий пан//]

Очистити

Фонетична транскрипція

Результат транскрибування

[п°о ве°с°н°ї 1663 р°ок°у дв°о°їє п°од°ор°ожн°їх/ верхї н°д°обрїх к°о°н°`°ах/ ізближализ°
 д°оки°їєва° з°б°їл°ог°ор°о°з°с°к°ог°о шл°`°ах°у// одїн б°уї м°о°л°одїї с°об°ї к°озак/
 збр°оїнїї/ ј°аг°д°ов°їїнїї/ др°угїї п°оодеж°ї іп°осив°її б°ор°о°д°ї/ сказа°д°би/ п°їп/ ап°о
 шабл°`°у°ц°ї п°їдр°`°ас°о°ї°у/ п°оп°їст°о°л°`°ах зап°о°ї°ас°о°м°`°їп°од°оїгїх шр°а°м°ах н°вїд°у/
 старїї к°озарл°`°уга// к°о°н°їїїїх п°от°о°млє°н°її/ одежа іт°ор°оки п°озапїл°`°ув°а°н°її/ зараз
 б°ул°о зн°аїти/ шч°о ід°у°т° нєзблиз°ка//

нєд°оїїз°д°`°ачи верст°ої з°одв°ї чїз°отри д°оки°їєва/ вз°`°али в°о°нї у°л°їв°у р°ук°у
 даїп°обралїз° г°аїїєм/ п°окрїв°її д°о°р°їз°ц°ї// іхт°о° т°їл°ки бачїї/ ј°акв°о°нї зп°о°л°`°а
 п°оверн°їїли їгаї/ у°с°`°аке зараз д°о°мї°с°л°`°ал°о°с°/ к°уди в°о°нї пр°ост°уїїу°т°// крїва
 д°о°р°їжка вєла д°очерєв°а°нєв°ог°о х°ут°ора/ хм°арїшча// ачерєв°а°н° б°уї т°`°ажк°о
 гр°ошовїтїї даївесєлїї п°а°н°//]

Очистити

Гриценко Я. Автоматичний генератор транскрипції для українськомовних текстів [Електронний ресурс] / Яна Гриценко – Режим доступу до ресурсу: <https://github.com/YanGrits/Phonet>.

Посилання містить файли:

- Transcription та Transcription2 – власне алгоритм для генерації транскрипції;
- main – сервер;
- static - клієнт для зв'язку та взаємодії з сервером, бібліотеки для створення інтерфейсу.