

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

## **СТВОРЕННЯ КОРПУСУ УСНОГО ДИТЯЧОГО МОВЛЕННЯ**

**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 ДУТЧАК**

**Науковий керівник:**  
Валентина РОБЕЙКО

**Рецензент:**  
канд. філол. наук Зоя ДУДНИК

**Київ – 2021**

# ЗМІСТ

<b>ВСТУП</b>	<b>2</b>
<b>РОЗДІЛ 1. СУЧАСНИЙ СТАН ДОСЛІДЖЕННЯ ТА ПРОБЛЕМИ СТВОРЕННЯ АКУСТИЧНИХ КОРПУСІВ ДИТЯЧОГО МОВЛЕННЯ</b>	<b>4</b>
1.1. Психолінгвістика дитячого мовлення	4
1.1.1. Мовлення в онтогенезі. Етапи формування мовлення в дитини	4
1.1.2. Проблематика дослідження дитячого мовлення	6
1.2. Корпуси усного дитячого мовлення	9
1.2.1. Корпус усного мовлення. Основні аспекти визначення	9
1.2.2. Акустичні корпуси та інструментарії для їх створення	11
1.2.3. Програми створення корпусів дитячого мовлення ELAN та PRAAT	14
<b>ВИСНОВОК ДО РОЗДІЛУ 1</b>	<b>16</b>
<b>РОЗДІЛ 2. СТВОРЕННЯ КОРПУСУ УСНОГО ДИТЯЧОГО МОВЛЕННЯ: ЗАПИС ЗВУКОВОГО МАТЕРІАЛУ Й АНАЛІЗ ФОНЕТИЧНИХ ОСОБЛИВОСТЕЙ ДИТЯЧОГО МОВЛЕННЯ</b>	<b>16</b>
2.1. Підбір дикторів і запис звукового матеріалу	16
2.2. Сегментація записів	17
2.3. Особливості анотування корпусу	18
2.4. Фонетичні особливості дитячого усного мовлення	21
2.4.1. Особливості формант голосних звуків	21
2.4.2. Характеристика інтенсивності дитячого мовлення	21
2.4.3. Особливості інтонаційних контурів дитячого мовлення	23
2.4.4. Частота основного тону дитячого мовлення	23
2.4.5. Тривалість синтагм дитячого мовлення	24
2.4.6. Характеристика дефектів мовлення у дітей	25
<b>ВИСНОВОК ДО РОЗДІЛУ 2</b>	<b>26</b>
<b>РОЗДІЛ 3. СТВОРЕННЯ ВЕБ-САЙТУ ДЛЯ КОРПУСУ УСНОГО ДИТЯЧОГО МОВЛЕННЯ</b>	<b>28</b>
3.1. Структура веб-сайту	28
3.2. Програмні засоби для створення сайту: HTML і PYTHON	29
3.3. Jinja2 – система веб-шаблонів для створення .html файлів	30
3.4. Додаткове файлове забезпечення: аудіозаписи, файли для завантаження, анотації записів у форматі .zip тощо	32
<b>ВИСНОВОК ДО РОЗДІЛУ 3</b>	<b>33</b>
<b>ВИСНОВОК</b>	<b>34</b>
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</b>	<b>36</b>
<b>ДОДАТКИ</b>	<b>41</b>
Додаток 1	41
Додаток 2	42
Додаток 3	45
Додаток 4	46
Додаток 5	49
Додаток 6	51
Додаток 7	52
Додаток 8	53

## **ВСТУП**

Попри те, що формування мовлення дорослої людини починається з раннього віку, феномен розвитку дитячого мовлення досі залишається не до кінця дослідженим. Тому зараз дослідження дитячого мовлення є однією з **актуальних**, центральних і найбільш важливих галузей лінгвістики й передбачає застосування комп'ютерних технологій, що забезпечують запис, розшифровку, кодування й автоматичний аналіз фонетики мовлення дитини.

Сьогодні одним із провідних підходів до вивчення мовних засобів є корпусний підхід, що не лише дає змогу аналізувати великий обсяг мовного матеріалу, а й розширює можливості наукового пошуку.

Спираючись на спостереження й дослідження фонетичних характеристик дитячого мовлення, можна зрозуміти, що всі дитячі відступи від норми завжди мають причину або закономірність, яка допомагає виявити, як влаштовані справжні мовні правила, де аномалії в мові, де закономірності, що логічно, а що – ні.

Дослідженням лінгвістики дитячого мовлення нині займаються такі відомі науковці й наукові організації, як К. Сноу та Б. Мак-Вінні, А. Шахнарович, Є. Кубрякова, Є. Ісеніна, Є. Протасова, В. Харченко, Міжнародна асоціація з вивчення дитячого мовлення (IASCL) тощо.

У роботі було узагальнено результати теоретичних і практичних досліджень фонетичних характеристик українського дитячого мовлення на базі створеного корпусу дитячого мовлення, а також розроблено веб-сайт із вільним доступом до створеного корпусу для всіх користувачів мережі Інтернет.

**Об'єктом** досліджень цієї роботи є усне мовлення дітей.

**Предмет** – його фонетичні характеристики. Дослідження виконане на **матеріалі** усномовних записів, що увійшли до створеного корпусу.

**Метою** кваліфікаційної роботи є створення й аналіз корпусу дитячого мовлення, а також розроблення веб-сайту для розміщення в вільному доступі цього корпусу й інформації з нашого дослідження, яка стосується як дитячого мовлення загалом, так і його окремих фонетичних характеристик. Оскільки

подібний корпус вперше викладається в мережу Інтернет для вільного користування, він може стати в нагоді студентам, викладачам, психологам, психолінгвістам, логопедам, педагогам, а також усім, хто цікавиться дослідженнями українського дитячого мовлення, що й становить **практичну цінність** і **новизну** цього дослідження. На сайті зібрана інформація, яка полегшить майбутні дослідження науковців для вдосконалення вже існуючого корпусу усного дитячого мовлення. На сайті також передбачено можливість розширювати й поповнювати корпус різними дослідженнями на тему усного дитячого мовлення й постійно вдосконалювати вже наявний корпус.

Реалізація мети кваліфікаційної роботи здійснюється за допомогою виконання таких **завдань**:

1. З'ясувати сучасний стан досліджень дитячого мовлення в Україні.
2. Зібрати матеріал для досліджень (звукові записи усного дитячого мовлення).
3. Дослідити особливості анотування й сегментації усного дитячого мовлення.
4. Дослідити фонетичні особливості дитячого усного мовлення.
5. Виконати узагальнений аналіз українського усного дитячого мовлення.
6. Розробити веб-сайт корпусу усного дитячого мовлення з можливістю користування ним у вільному доступі.

У кваліфікаційній роботі використано такі **методи** дослідження: описовий метод, метод аналізу й синтезу, експериментально-фонетичний метод.

Робота складається зі вступу, трьох розділів, висновків і списку використаної літератури.

## РОЗДІЛ 1. СУЧАСНИЙ СТАН ДОСЛІДЖЕННЯ ТА ПРОБЛЕМИ СТВОРЕННЯ АКУСТИЧНИХ КОРПУСІВ ДИТЯЧОГО МОВЛЕННЯ

У цьому розділі ми розглянемо психолінгвістику дитячого мовлення та етапи розвитку мовлення у дитини. Більш детально дізнаємось про проблематику досліджень усного дитячого мовлення. З'ясуємо, що таке корпус усного дитячого мовлення, які корпуси існують у світі та якими програмами користуються дослідники, щоб створити корпус.

### 1.1. Психолінгвістика дитячого мовлення

#### 1.1.1. Мовлення в онтогенезі. Етапи формування мовлення в дитини

Щоб своєчасно й правильно діагностувати відхилення та грамотно побудувати корекційно-виховну роботу задля лікування мовних патологій, необхідно розуміти закономірності дитячого мовленнєвого розвитку.

Згідно з дослідженням О. Воротинцевої «Основні етапи розвитку мовлення в дітей» [6], існують такі етапи мовленнєвого розвитку в дитини:

#### 1. 3 місяці – 2 роки.

- **3 місяці:** дитина чує голоси й розпізнає вирази обличчя дорослих. Починає відгукуватись на голос. Згодом за допомогою звуків малюк починає виражати емоції.
- **6 місяців:** вимовляє більшу частину звуків і склади слів. Передає свій настрій інтонацією, відгукується на своє ім'я. Інколи в цей період дитина може вимовляти такі прості слова, як «мама» і «тато», але вона не вкладає в них ніякого змісту, а просто повторює випадкові склади.
- **9 місяців:** розуміє такі слова, як «привіт», «ні», «так». Використовує інтонації й більшість приголосних.
- **12 місяців:** вимовляє й усвідомлює значення слів «мама» і «тато». Виконує найпростіші вказівки: «Не можна».
- **18 місяців:** дитина в цьому віці вимовляє прості слова та вказує на предмети, які їй називають. Може повторювати за дорослим останні

слова або склади в реченні, при цьому пропускаючи перші склади (наприклад, «люнок» замість «малюнок»).

## **2. 2-3 роки.**

- До кінця 3-го року життя словниковий запас дитини складає 500 слів.
- Розуміє й виконує прості вказівки: «*Піди до столу й візьми ложку*».
- Вимовляє більшість приголосних, окрім сонорних [л] і [р].
- Засвоює мовлення в діалозі.
- Засвоює такі поняття, як: «багато» й «один».

## **3. 3-4 роки.**

- У чотирирічному віці словниковий запас мовлення дитини стрімко зростає до 1500 слів.
- Діти повністю засвоюють рідну мову.
- Починають будувати складні речення.
- Знають, як їх звати, свою стать і можуть відповісти на питання «*Скільки тобі років?*»
- У лексиконі дитини з'являються такі нові слова, як: «чому» й «навіщо».

## **4. 4-5 років.**

- Дитина вміє будувати складносурядні й складнопідрядні речення.
- Словниковий запас складає 1500-2000 слів.
- Малюк іще вчиться правильно використовувати відмінкові закінчення й узгоджувати слова.
- Може переповідати казки й розповідати невеликі віршики.
- Може висловити власну думку з приводу певних подій і явищ.

## **5. 5-6 років.**

- Словниковий запас становить 3500 і більше слів.
- Дитина використовує всі частини мови в мовленні.
- Опановує всі звуки.
- Розвиває мовну й мовленнєву увагу та логічне мислення.

- Використовує антоніми й синоніми.
- У віці 5 років дитина легко переказує оповідання, що свідчить про здобутки в засвоєнні досить складного виду мовлення – монологічного.

Як бачимо, жоден з описаних нами етапів не характеризується чіткими межами. Кожен попередній рівень плавно переходить у наступний. Початковий етап оволодіння мовою людина зустрічає в процесі життєво необхідної потреби спілкування. Мова дитини розвивається не в монологі, а в діалозі – у комунікації з дорослими. Найбільш інтенсивно мова розвивається саме в ранньому віці (сензитивний період розвитку мови – 2-6 років). Це пояснюється стрімким процесом дозрівання й пластичністю ЦНС у цей період. Якщо цей період прогавити, тобто позбавити дитину мовної стимуляції внаслідок сенсорної (при порушеннях слуху) або соціальної депривації, будуть спостерігатися затримки в розвитку мови або повна її відсутність.

### **1.1.2. Проблематика дослідження дитячого мовлення**

Дитяче мовлення є особливим етапом онтогенетичного розвитку мовлення людини. Це дивовижний феномен, до вивчення якого свого часу закликали такі дослідники, як П. Гальперін, О. Леонтьєв, С. Рубінштейн [12].

Дослідження мовлення дітей завжди викликало значний інтерес у психолінгвістів, бо воно змушує розширити наші уявлення не тільки про оволодіння мовою, а й про саму мову, про те, як вона функціонує в мовленнєвій діяльності, не тільки про становлення пізнавальних та інших психологічних процесів, але й про їх побудову, а також про таке поняття, як «людина» взагалі. Більше того, дитяче мовлення традиційно вважається таким самим природним об'єктом вивчення для психолінгвістики, як, наприклад, закони словотвору. Проблематика, пов'язана з дитячим мовленням, включається в лекційні курси й програми іспитів за фахом «психолінгвістика». Проте досі не існує чітко окреслених розроблених концепцій, які стосувалися б винятково дитячого мовлення. Тобто не існує теорії, яка пояснювала б, як насправді формується

мова дитини, як оволодіння мовою стає базою розвитку в дитини всієї сукупності пізнавальних процесів – сприйняття, узагальнення, визначення мети, оцінки.

Оскільки онтолінгвістику відносять до розряду сфер дослідження, які на даний момент і надалі формуються, то її не можна вивчати без звернення до дитячої мовної свідомості, яка пов'язана з засвоєнням мови як одного з основних інструментів пізнання світу.

Адже пройти весь шлях до здобуття такої здатності, як говорити, дитина не може, не пізнавши певних важливих особливостей формування мови. Тому дорослі й не розмовляють лише простими реченнями, а обов'язково вносять корінь слова в контекст, варіюючи з фразами, задаючи валентність словам, періодично виділяючи компоненти: *«Дивись, котик, ой, який котик, зараз ми його погодуємо»*. Якщо спостерігати за діалогом дитини й дорослого, можна виявити, що другий частіше повторює за першим його стиль мовлення, а не навпаки, показуючи цим приклад спілкування [7].

*«Дослідження дитячого мовлення набули актуальності, коли дитину почали сприймати інакше, ніж дорослого, а аналіз розвитку мовленнєвої діяльності в онтогенезі почали розглядати, спираючись на твердження, що дитяче мовлення – це не лише етап у формуванні дорослого мовлення, а особлива система з певним набором правил, що варіюються від одного періоду до іншого»* [2, с. 192].

Засвоєння дитиною звукової сторони мовлення, граматики, лексики, тобто «додаткових компонентів», і їхня інтеграція дали змогу вивчати феномен «дитячого мовлення». Він усе більше привертає увагу не тільки лінгвістів, а й психологів, логопедів, дослідників розвитку мовлення й багатьох інших фахівців [3].

Серед найактуальніших сучасних зарубіжних досліджень на тему дитячого мовлення можемо виділити:

1. Бази даних аудіозаписів дітей із нетиповим розвитком мовлення: *A new method for collection and annotation of speech data of atypically developing children* [19].
2. Розпізнавання мовлення дитини у взаємодії «людина-робот»: *Advances in Automatic Speech Recognition for Child Speech Using Factored Time Delay Neural Network* [33].

Також велика увага приділяється дослідженням порушень дитячого мовлення. Серед найпопулярніших робіт на цю тему, можемо виділити:

1. **Speech Perception Skills of Children With Speech Sound Disorders: A Systematic Review and Meta-Analysis**[39]. Дослідження провели, щоб виявити, чи є у дітей дошкільного та раннього шкільного віку з порушеннями звукового мовлення (ССД) труднощі із сприйняттям мови.
2. **What Do Children with Speech Sound Disorders Think about Their Talking** [44]? Стаття спрямована дослідити, чи помічають діти з мовними розладами різницю між їх мовленням та мовленням дорослих, і що вони відчувають у момент розмови.
3. **Reconceptualizing practice with multilingual children with speech sound disorders: people, practicalities and policy**[37]. Професії логопедів та мовних терапевтів необхідні для надання послуг дедалі більшій кількості дітей з білінгвальних сімей. Дане дослідження зосереджене на проблемі практики логопедів та мовних терапевтів з дітьми, що говорять декількома мовами.

Сьогодні для української лінгвістики також є особливо актуальним дослідження дитячого мовлення, оскільки:

- по-перше, зарубіжні науковці володіють великою теоретичною й практичною базою поетапного й глибокого вивчення мовлення в онтогенезі;

- по-друге, щоб досліджувати цю проблему, українські вчені брали за матеріал своїх досліджень не українську мову, а російську, англійську або французьку [17].

Сьогодні такі українські мовознавці, як О. Моїсеєнко, О. Снітко, І. Сухан, Н. Зозуля, Н. Дзюбишина-Мельник та О. Рульова, займаються питаннями діяльності в онтогенезі мовлення [17].

Інновація дитячого мовлення – унікальний феномен, який дозволяє побачити приховані механізми лінгвокреативного мислення в дії, що дадуть змогу в майбутньому відкрити специфіку погляду дитини на світ і на мову.

## **1.2. Корпуси усного дитячого мовлення**

### **1.2.1. Корпус усного мовлення. Основні аспекти визначення**

У наукових дослідженнях різних мов світу часто фігурує таке поняття, як «корпус». Лінгвістичний термін «корпус» дорівнює поняттю «електронний корпус» і має обов'язково підходити для машинного читання.

Також при дослідженні певної мови ми спираємося на «корпуси текстів» та «корпуси усного мовлення».

Корпус текстів – це стандартно організована, програмно опрацьована, репрезентативна вибірка текстів певної мови, писемних, а також затранскрибованих усних текстів без відповідних аудіозаписів, що призначена для наукового й практичного вивчення.

Корпус усного мовлення – велика колекція аудіозаписів розмовної мови. Більшість усних корпусів також містять додаткові текстові файли, що містять транскрипції вимовлених слів і час, коли кожне слово відбувалося в записі [38].

Ми можемо стверджувати, що усна форма мовлення безумовно є основною, фундаментальною, вихідною. Наші уявлення про будь-яку природну мовою є неповними й навіть спотвореними, якщо вони не ґрунтуються на даних усної форми цієї мови. Корпусні лінгвістичні дослідження неодмінно мають включати дослідження корпусів усного мовлення як один із центральних

компонентів. Такі корпуси існують для тих мов, на матеріалі яких корпусні дослідження досить розвинені [11]. Наприклад, для корпусу англійської мови.

Аналізувати усне мовлення, коли воно існує тільки у формі звуку, практично неможливо: звук має занадто багато аспектів, щоб людський погляд і свідомість могли одночасно утримати якийсь його значимий фрагмент. Тому необхідним є спосіб об'єктивної фіксації усного мовлення. Імовірно, найбільш зручний спосіб такої фіксації – переведення її у графічний вигляд, тобто створення транскрипції. Корпус усного мовлення має складатися щонайменше з двох компонентів – аудіозаписів мови й відповідних для записів текстових транскрибованих файлів [24].

Також ми можемо поділити корпуси усного мовлення ще на 2 групи [8]:

1. Усне читане мовлення. Наведемо приклади:

- трансляції новин: *1997 English Broadcast News Speech (HUB4)* [28];
- списки слів: *BNC Frequency lists from the companion web site to the book: Leech, Geoff, Paul Rayson & Andrew Wilson (2001)* [45];
- цифрові послідовності: *Comprehending Real Numbers: Development of Bengali Real Number Speech Corpus* [25].

2. Спонтанне мовлення. Наведемо приклади:

- діалоги й зустрічі – жива розмова двох або більше людей: *The ICSI Meeting Corpus* [40];
- розповіді – людина, що розповідає історію: *Estonian North Wind and the Sun Corpus v.1.0.3* [30];
- усне пояснення маршруту: *Aalto University DSP Course Conversation Corpus 2013-2016* [42];
- усне пояснення завдання: *Hamburg Modern Times Corpus* [30];
- моделювання різноманітних життєвих ситуацій, наприклад, бронювання білету: *CLAPI* [20];
- парламентські засідання: *Plenary Sessions of the Parliament of Finland* [31].

За цими типами існують:

- корпуси спонтанного мовлення: *Route to A wing* [34];
- корпуси лекцій і семінарів: *ELFA Corpus* [27];
- національні корпуси мов: *Spoken English Corpus (SEC)* [43];
- корпуси текстових статей: *The Spoken Wikipedia Corpora* [26];
- корпуси літературної мови: *Corpus for the Study of Contemporary French* [22];
- корпуси усного дитячого мовлення: *CHILDES* [35] і т.д.

### 1.2.2. Акустичні корпуси та інструментарії для їх створення

Створення акустичних корпусів нині є самостійним і популярним напрямом мовленнєвих технологій. Проте це доволі складна технологічна задача, яка потребує значних фінансових і кадрових ресурсів. Для її вирішення в 1990-х рр. у світі були створені координаційні центри для збору, зберігання, розповсюдження та створення загальнодоступних і стандартизованих мовних ресурсів, зокрема мовленнєвих [16].

Сьогодні існує безліч корпусів зокрема дитячого усного мовлення. Дослідження в цій галузі проводяться в усьому світі й зараз викликають неабиякий інтерес в українських лінгвістів. Нижче подано приклади деяких корпусів усного дитячого мовлення.

#### **BiRCh**

BiRCh – *Corpus of bilingual russian child speech* [8].

Мета проєкту – створення розмірного, синхронізованого й вільно доступного в Інтернеті корпусу мовлення дітей, які вивчають російську мову в монолінгвальних і білінгвальних сім'ях.

Хоча мовлення мігрантів часто піддається негативній стереотипізації (навіть самими мігрантами), воно відіграє ключову роль у культурній ідентичності й практиці російськомовних громад, а також дає можливість виокремити основні фундаментальні властивості мови, процесів засвоєння й розвитку мовлення, що відбуваються в житті людини. Граматично розмічений корпус монолінгвального й білінгвального дитячого мовлення надасть необхідні

дані для досліджень мови в середовищі імігрантів і монолінгвальних громадах, дані про розвиток херітажних мов, розвиток мовлення в цілому й видозміни мовлення в ситуації білінгвізму й мовного контакту.

Корпус також надасть необхідну інформацію для практичних цілей – розробки мовних матеріалів для мігрантів другого покоління, які вивчають російську мову; для батьків, які виховують білінгвальних дітей; а також для фахівців-практиків, розробників навчальних програм та експертів у галузі мовної політики при розробці нормативів і стратегій.

Автори:

1. Е.В.Денисова-Шмидт, доцент (*Lecturer*), Школа гуманітарних і соціальних наук, Університет Санкт-Галлена, Швейцарія.
2. И.Е. Дубініна, доцент (*Assistant Professor*) російської мови, директор програми з вивчення російської мови, Університет ім. Брандейса.
3. С. А. Маламуд, професор (*Associate Professor*) лінгвістики, Університет ім. Брандейса.

### **CHILDES [35]**

CHILDES – міжнародна система обміну даних дитячого мовлення (*Children Language Database Exchange System*), розроблена американськими вченими К. Сноу та Б. Маквіни в Пітсбургському університеті, яку використовують для аналізу розмовного мовлення, спонтанного мовлення дітей і їхніх батьків, а також для дослідження засвоєння другої мови [35]. Нині база даних CHILDES – корпус значного обсягу, який містить тексти мовами різноманітного типологічного рангу, об'ємну бібліографію зі психолінгвістики, теорії засвоєння першої та другої мов, а також правила введення матеріалів і пакет програм для їхнього аналізу. Кожен дослідник може скористатися необхідними йому даними, а також розмістити в CHILDES свої матеріали. Система має широку сферу застосування, і її можна конвертувати в інші лінгвістичні системи.

Один із блоків бази даних CHILDES складають усні розповіді за серією картинок «*Frog where are you?*» – історії пригод хлопчика і його собаки, які шукають жабку. Такий блок представлений у CHILDES не випадково: здатність дитини побудувати розгорнуту розповідь про якусь подію й основні параметри такої розповіді є одними з основних критеріїв для визначення рівня розвитку мовлення. У системи є такі переваги, як доступність, різноаспектність, єдиний формат записів, різноманітність матеріалів і їхня достовірність [9].

### **The CMU Kids Corpus [29]**

Ця база даних представлена читанням різноманітних речень дітьми вголос. Спочатку вона була створена для того, щоб розробити навчальний набір із голосів дитячого мовлення для автоматичного розпізнавання мовлення SPHINX II і його використання в проєкті LISTEN в Університеті Карнегі-Меллона. Вікова категорія мовців складала 6-11 років, тобто це були діти, які навчались у 2-6 класах. Було записано понад 5 тисяч висловлювань 24-ох чоловічих і 52-ох жіночих голосів. Диктори походили з двох окремих груп населення. Було обрано мовців без значних дефектів мовлення з досить чистою вимовою. Система LISTEN також потребувала прикладів помилкового читання й діалектних варіантів вимови. Тому діти, яким був властивий цей тип мовлення та які походили зі шкіл з високим ризиком поганого засвоєння навички читання, могли отримати вигоду від участі в цьому проєкті, отримавши консультацію від будь-якого викладача або системи, побудованої на цій базі даних. Цю підгрупу назвали FP. У цій групі налічувалося 32 мовця й понад 1800 їхніх висловлювань. Список мовців, група, в якій вони перебувають, і кількість речень на одного диктора можна знайти в каталозі «*tables*» у файлі під назвою «*speaker.tbl*».

Крім наведених вище корпусів усного дитячого мовлення, існує низка подібних систем. Серед них *CSLU Kids' Speech Corpus*, *PF-STAR Children's Speech Corpus*, *CALL-SLT*, *Providence Corpus*, *Swedish NICE Corpus*, *SingaKids-Mandarin*, *CHIEDE*, *TIDIGITS*, *Рассказы о сновидениях*, *Demuth Sesotho Corpus* тощо.

Створення корпусів усного дитячого мовлення є важливим аспектом у вивченні формування мовлення людини. Мовленнєвий матеріал, що міститься в корпусі дитячого мовлення, використовується при проведенні міждисциплінарних досліджень різних аспектів становлення мовлення й їхнього зв'язку з когнітивним та емоційним розвитком дитини. Корпус усного дитячого мовлення допомагає нам почути, затранскрибувати й дослідити приховані механізми людського мислення в дії, що дає змогу в майбутньому відкрити специфіку й відстежити прогресію поглядів дитини на світ і на мову.

### **1.2.3. Програми створення корпусів дитячого мовлення ELAN та PRAAT**

ELAN – відкритий безкоштовний програмний ресурс, створений нідерландським Інститутом психолінгвістики імені Макса Планка (*The Language Archive*, м. Неймеген) [14].

Програма ELAN вважається однією з найкращих для створення акустичних корпусів – здебільшого через зручний функціонал, безкоштовність і наявність інструкції *Help*, яка чітко пояснює особливості роботи з програмою.

ELAN – це перш за все спільне представлення звукової та графічної інформації, що дає змогу здійснювати багаторазові лінгвістичні розмітки. Так, наприклад, якщо у вживаних проектах транскрибуються такі явища, як паузація, рух тонової кривої в акцентованих словоформах і тип ілюзійно-фазового значення ЕДЕ, то вони кодуються в рамках одного рядка, а саме в аналогічній формі. Кожен документ ELAN, створений для анотування усного українського мовлення, складається з двох файлів: одного медіафайлу й одного анотаційного файлу. Медіафайл – це файл звукозапису у *.wav*-форматі, анотаційний файл – створений у програмі файл ELAN із розширенням *.eaf*.

Особливості транскрипції усного мовлення в програмі ELAN такі: по-перше, потрібно спиратися на акустичний сигнал і виявляти специфіку вимови дикторів; по-друге, потрібно дотримуватися протоколу використання

кириличних і IPA-символів, що дозволяє узгодити корпус, створений на базі програми ELAN із міжнародними фонетичними стандартами.

За допомогою програми ELAN досить легко здійснювати сегментацію й анування звукових файлів, особливо при створенні корпусу усного мовлення, а потім описувати отриманий обсяг інформації.

*Praat* – голландський безкоштовний програмний ресурс для фонетики, який може записувати й аналізувати мовленнєві звуки в *Mono* та *Stereo* форматах. *Praat* розроблений для лінгвістів і студентів мовознавства, а також інших вчених і дослідників, яким потрібно аналізувати людську мову. Ресурс можна знайти у вільному доступі для будь-якого користувача. Програма підтримує синтез мовлення, зокрема також артикуляційний синтез. Вона також створює якісні зображення даних, які можуть бути опубліковані [21].

Інтерфейс програми складається з двох частин: *Praat Objects* і *Praat Picture*. Також вона має інструкцію *Help*, яка чітко пояснює особливості роботи з програмою. За допомогою *Praat* можна записувати аудіофайли або імпортувати й аналізувати вже записані. Програма складається з чотирьох основних видів аналізу: спектральний, тоновий, формантний та аналіз інтенсивності. Також можна відредагувати аудіозаписи так, як необхідно для певного дослідження, щоб заощадити час, а також можна створювати скрипти для автоматизації роботи [36].

Здійснювати за допомогою *Praat* деякі функції аналізу зручніше, ніж за допомогою ELAN. Тому в цій кваліфікаційній роботі використовували обидва програмні засоби. ELAN може працювати з голосовими формами хвиль, але не має спектрального аналізу й функцій відображення тону, необхідних для голосового аналізу. *Praat* добре відомий у лінгвістиці як безкоштовне програмне забезпечення для аналізу голосу. *Praat* дозволяє вказати конкретні частини аудіофайлу для детального аналізу, включаючи спектральний аналіз. Для того, щоб використовувати *Praat* разом з ELAN, необхідно використовувати окреме програмне забезпечення під назвою *sendpraat*, яке передає аудіо з ELAN у *Praat* [41].

## ВИСНОВКИ ДО РОЗДІЛУ 1

У цьому розділі ми розглянули психолінгвістику дитячого мовлення, дослідили основні етапи формування мовлення у дитини та виявили, що дитяче мовлення розвивається не в монолозі, а в діалозі.

Також ми окреслили проблематику досліджень феномену дитячого мовлення, описали найпопулярніші дослідницькі роботи на цю тему серед українських та зарубіжних лінгвістів, і виявили, що переважно українські вчені брали за матеріал своїх досліджень не українську мову, а російську, англійську або французьку, тому вітчизняна лінгвістика зараз неабияк потребує нового погляду на дану проблему.

Далі ми визначили, які існують акустичні корпуси і в чому їх особливості. Та на завершення, ми дослідили дві найпопулярніші програми для створення корпусу ELAN та PRAAT і зробили висновок, що обидві програми в комплексі є досить зручними у створенні корпусів усного мовлення.

## **РОЗДІЛ 2. СТВОРЕННЯ КОРПУСУ УСНОГО ДИТЯЧОГО МОВЛЕННЯ: ЗАПИС ЗВУКОВОГО МАТЕРІАЛУ Й АНАЛІЗ ФОНЕТИЧНИХ ОСОБЛИВОСТЕЙ ДИТЯЧОГО МОВЛЕННЯ**

*У цьому розділі ми познайомимось з усіма етапами створення даного корпусу: від підбору дикторів та запису звукового матеріалу, до процесу анотування, сегментації та досліджень фонетичних особливостей усного дитячого мовлення. За допомогою наших досліджень, ми зможемо зробити узагальнений аналіз особливостей усного дитячого мовлення.*

### **2.1. Підбір дикторів і запис звукового матеріалу**

У попередньому дослідженні ми підібрали шість дикторів віком 5-6 років. Цей вік вважається найбільш оптимальним для досліджень дитячого мовлення. У цей період у дитини активно розвивається навичка слухового контролю власної вимови, уміння виправляти себе й опанувати один із найскладніших

видів мовлення – монологічне. Діти такого віку мають досить багато дефектів вимови, які є однаковими для всіх представників цієї вікової категорії, але вже формують індивідуальний стиль мовлення.

Запис мовлення кожного диктора в середньому тривав близько 10 хвилин. Ми заздалегідь підготували запитання на різні теми, на які діти мали дати розгорнуту відповідь. Далі ми здійснили синтагматичне членування мовного потоку: розчленували думки на окремі, певним чином пов'язані між собою смислові комплекси – синтагми. Членування звукового запису відбувалося з урахуванням меж синтагм за паузами, індивідуальними особливостями вимови диктора (дефекти мовлення, дихання, наявність подовження голосних і приголосних, довгі паузи між синтагмами тощо) і контексту.

Для анотації ми використовували спеціальні позначки, розроблені заздалегідь, щоб проілюструвати всі особливості мовлення дикторів. Наприклад: «\_ем\_» – довге [ем], «\_запереч\_у» – заперечний звук [у], «\_каш\_» – кашель. Далі ми виявили й описали особливості індивідуального мовлення дикторів та охарактеризували проанотовані файли (див. Додаток 4).

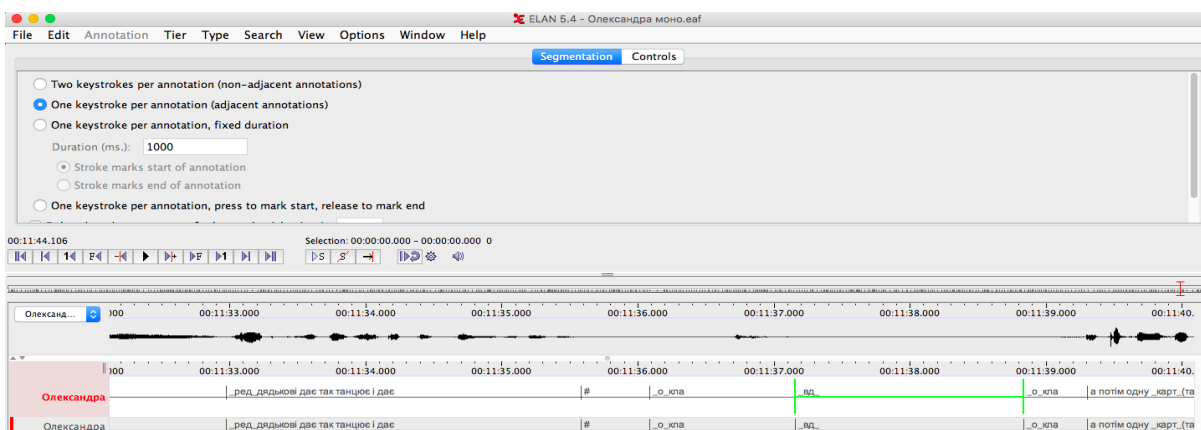
## 2.2. Сегментація записів

*«Синтагма – це одиниця, яка вичленовується із фрази відповідно до змісту висловлювання. Наприклад, вислів **Допомогти не можна відмовити**, якщо його не поділити на синтагми, буде позбавлений змісту. Синтагма виконує дві функції: фонетичну й семантичну» [10, с. 71].*

Членування звукового запису відбувалося з урахуванням таких особливостей:

1. основним критерієм було визначення межі синтагми за паузами. Дитяче живе мовлення досить важко ділити на чіткі, інтонаційно завершені синтагми, тому ми вирішили вважати синтагмою відрізок мовлення між двома фізично вираженими паузами, або якщо пауза була психологічною, або коли мовлення дикторів накладається одне на одне;

2. визначили індивідуальні особливості вимови диктора (дефекти мовлення, дихання, наявність подовження голосних і приголосних, довгі паузи між синтагмами тощо);
3. застосовували семантичний критерій (враховували контекст), тобто розглядали синтагму як одиницю, що виражала в певному контексті одне поняття [1].



Малюнок 2.1. Приклад сегментування запису в програмі ELAN (див. Додаток 1).

### 2.3. Особливості анотування корпусу

Одна синтагма мовця у звуковому записі тривала в середньому 0.5-1.00 с. Для анотації ми розробили власну систему позначень (див. Додаток 4).

Можемо вивести такі характерні особливості мовлення кожного диктора:

Характерно	Приклад	Час
Розтягування звуків	_роз_(передостанній день весни)	00:08:00-00:10:00
Голосний вдих/видих	_вд_	00:51:40-00:51:80
Вимова по складах	_пскл_настя	03:23:00-03:23:50

Таблиця 2.1.

**АНЯ**  
(Аня 5 років елан2.eaf)

<b>Характерно</b>	<b>Приклад</b>	<b>Час</b>
Картавість	а потім одну _карт_(тарілку розбила)	11:40:00-11:41:00
Звуки _и_, _і_	_и_ _і_	10:32:50-10:33:50 07:35:50-07:36:50
Нечітка вимова	_нечіт_навіть	03:42:09-03:42:80

**Таблиця 2.2.**

**ОЛЕКСАНДРА**  
(Олександра моно.eaf)

<b>Характерно</b>	<b>Приклад</b>	<b>Час</b>
Картавість	_с_русский поет	03:44:50-03:45:50
Схлипування	_схл_	03:19:60-03:20:00
Редукція	ну _ред_ незнаю	03:19:00-03:19:80

**Таблиця 2.3.**

**БОГДАН**  
(Богдан моно.eaf)

<b>Характерно</b>	<b>Приклад</b>	<b>Час</b>
Картавість	_карт_річка дала синій	10:07:00-10:08:00
Голосний вдих/видих	_вд_	10:10:00-10:10:50
Обрив слів	_о_в	02:40:60-02:40:90

**Таблиця 2.4.**

**ВЛАД**  
(Влад моно.eaf)

<b>Характерно</b>	<b>Приклад</b>	<b>Час</b>
Картавість	я їхав _карт_(на чорне море)	00:15:00-00:15:00
Звук _м_ (на початку	_м_ деколи машинками	00:29:00-00:30:00



## **2.4. Фонетичні особливості дитячого усного мовлення**

### **2.4.1. Особливості формант голосних звуків**

*«Звуки мови відрізняються один від одного насамперед набором обертонів. Обертони, з яких складається певний звук мови, називаються формантами. Визначальними в розпізнаванні голосних звуків є перші дві форманти. Наприклад, за деякими даними, для а – це приблизно 700 і 1200 Гц, для о – 400 і 800 Гц, для у – 300 і 700 Гц, для і – 200 і 2200 Гц, для и – 300 і 1900 Гц, для е – це 400 і 1600 Гц (у вимові різних людей висота формант неоднакова)» [10, с. 61].*

Щоб визначити середнє значення першої та другої форманти голосних дитячого мовлення, ми проаналізували декілька варіантів наголошених голосних і вивели з нього середнє значення. Далі ми занесли це в окрему таблицю для кожного диктора, а потім із них вивели загальне середнє значення формант дитячого мовлення (див. Додаток 5).

Проаналізувавши дані таблиць, ми можемо зробити такі висновки:

1. Найрідше в дитячому мовленні, представленому в нашому корпусі, використовуються наголошені [е] та [у].
2. Найвищі форманти в голосних [е] та [і].
3. Найнижчі форманти в голосних [о] та [у].
4. Форманти всіх голосних дитячого мовлення вищі, ніж у дорослих.
5. У формант дитячого мовлення немає істотних відмінностей залежно від статі (на протипагу дорослим мовцям).
6. Усі звуки голосних дитячого мовлення є дифузними.

### **2.4.2. Характеристика інтенсивності дитячого мовлення**

Інтенсивність мовлення – це густина потоку звукової енергії. Найменша інтенсивність звуку, яку ще може сприймати вухо людини (поріг чутності), становить 10-16 Вт/см<sup>2</sup>.

Кількісно інтенсивність звуку рівна середньому за часом потоку звукової енергії через одиничну площу, розташовану перпендикулярно напрямку поширення звуку.

На окремій синтагмі звукового запису кожного диктора, ми визначали мінімальне, максимальне й середнє значення інтенсивності мовлення. Для більш точного аналізу ми обрали десять синтагм. Далі для обчислення середнього значення інтенсивності ми вивели таку формулу:

$$(a1+a2+a3+a4+a5+a6+a7+a8+a9+a10)/10 = a \text{ (середнє)},$$

де  $a$  – будь-яке значення інтенсивності на відрізку амплітуди коливання (див. Додаток 6).

Проаналізувавши таблицю, подану в Додатку 6, ми зробили такі висновки щодо інтенсивності дитячого мовлення:

1. Мінімальне значення інтенсивності в диктора Богдана, а максимальне – у диктора Олександри.
2. Немає істотної різниці в показнику інтенсивності залежно від статі дитини.
3. Середнє значення середнього значення інтенсивності всіх дикторів становить 69 дБ.
4. Зазвичай, після заповнених пауз хезитації (у вигляді розтягнення вимови звуків), інтенсивність зростає (диктор Аня: «*\_у\_ роз\_голубого*»).
5. Під час самого розтягнення вимови звуку інтенсивність, зазвичай, досягає максимуму (диктор Злата: «*\_у\_*» – 80дБ).
6. Контур інтенсивності спадає при переліку певних предметів, ознак тощо (диктор Аня: «*рожеві жовті і голубі*»).
7. Інтенсивність звуку також зростає, коли дитина виділяє певне слово логічним наголосом (диктор Богдан: «*в школу радості*»).
8. Найчастіше на початку запису середній показник інтенсивності високий, а потім поступово спадає.

### **2.4.3. Особливості інтонаційних контурів дитячого мовлення**

«Інтонація – єдність взаємопов'язаних компонентів: мелодики, інтенсивності, тривалості, темпу мови і тембру вимови. Деякі дослідники вносять до складу компонентів інтонації паузи. Разом з наголосом інтонація утворює просодичну систему мови. Інтонація є важливим засобом формування висловлювання і виявлення його сенсу» [5].

Для того, щоб дослідити інтонаційні контури дитячого мовлення, ми обрали по декілька трислівних синтагм у записі кожного диктора з різними видами інтонації. Далі ми сформуваємо з даних графіки, які відтворюють особливості інтонаційних конструкцій кожної дитини (див. Додаток 2).

Проаналізувавши дані графіків, можемо сформулювати такі висновки:

1. Найчастіше в дикторів зустрічались такі види інтонацій, як рівномірно спадне завершення та рівномірно спадне незавершення.
2. У дівчат зустрічається більше видів інтонаційних контурів, ніж у хлопців.
3. Здебільшого діти використовують інтонації незавершення.

### **2.4.4. Частота основного тону дитячого мовлення**

Частота основного тону, яку ще позначають як ЧОТ, *Fundamental Frequency* або  $F_0$  – це частота коливання голосових зв'язок при вимові тонованих звуків (*voiced*). Коли ми вимовляємо нетоновані звуки (*unvoiced*), наприклад, якщо говоримо пошепки або вимовляємо глухі звуки, голосові зв'язки не коливаються, а значить ця характеристика для них не релевантна.

Частота основного тону може варіюватися в досить великих діапазонах, причому це може залежати не лише від людської статі (для нижчих чоловічих голосів у середньому частота становить 70-200 Гц, а для жіночих може досягати 450-500 Гц), але й бути характерним для однієї людини, особливо при емоційному мовленні. Нашим завданням було визначити середнє значення частоти основного тону для дитячого мовлення.

Ми проаналізували десять синтагм кожного диктора й визначили мінімальне, максимальне й середнє значення ЧОТ (див. Додаток 7).

Проаналізувавши дані таблиці, можемо зробити такі висновки:

1. Частота основного тону в мінімальних показниках у дівчат менша, ніж у хлопців.
2. Максимальне значення в усіх дітей знаходиться приблизно в однакових показниках.
3. Немає істотної різниці середнього значення ЧОТ у дітей залежно від статі (з віком це зміниться).
4. Мінімального значення ЧОТ досягає, зазвичай, при переліку певних предметів або коли дитині не зовсім цікаво брати участь у розмові (рівень емоційності значно нижчий, ніж зазвичай).

#### 2.4.5. Тривалість синтагм дитячого мовлення

Нагадаємо, що *“синтагма – це одиниця, яка вичленовується зі фрази відповідно до змісту висловлювання”* [10, с. 71].

Для аналізу тривалості синтагм дитячого мовлення, ми обрали такі типи інтонаційно завершеного мовлення:

1. розтягнута вимова звуків;
2. синтагми, що складаються з менше, ніж двох слів;
3. двослівні синтагми;
4. трислівні синтагми;
5. синтагми, що складаються з чотирьох і більше слів.

При дослідженні дитячого мовлення отримали такі результати:

Мовець	Тривалість запису	Кількість синтагм (загальна)	Розтягнута вимова звуків	Менше двох слів	Двослівні	Трислівні	Чотири і більше слів
Аня	10:20	135	16	34	21	20	44
Олекса ндра	11:59	204	15	56	38	27	68

Богдан	10:57	156	12	79	20	14	31
Злата	9:34	162	6	42	37	29	48
Данило	12:08	135	6	64	28	18	19
Влад	10:41	193	7	77	45	42	22

Таблиця 2.7.

#### КІЛЬКІСТЬ СИНТАГМ ВІДПОВІДНОЇ ДОВЖИНИ ДЛЯ КОЖНОГО ДИКТОРА

Із наявних результатів можемо зробити такі висновки:

1. Середня кількість синтагм, яку вимовляє дитина в віці 5-6 років протягом 10-12 хв, – 140 синтагм.
2. Зазвичай, діти найбільше проговорюють однослівні синтагми, а найменше – трислівні.
3. Хлопці вимовляють синтагми з більше ніж чотирьох слів набагато рідше, ніж дівчата (зазвичай, у віршованій формі (диктор Данило): «*і барвисті парасольки під дощем немов грибки*»).
4. Найбільш поширена однослівна синтагма, притаманна усім дикторам, – слово «*та*» (діалектизм, скорочена форма від «*так*»).
5. Немає істотної різниці вимовляння певних синтагм у дітей залежно від статі.
6. Деяким дітям притаманне хезитативне розтягнення певних звуків, яке може виокремлюватися в окрему квазісинтагму.

#### 2.4.6. Характеристика дефектів мовлення у дітей

Загальна інформація про особливості мовлення дітей 5-6 річного віку:

- Словниковий запас дитини складає 2500-3000 слів.
- У віці 5-6 років дитина повністю засвоює складову структуру складного слова: не плутає й не пропускає склади у слові.
- Особливо поліпшується вимова свистячих приголосних [с], [з], [ц] і

шиплячих [ш], [ж], [ч], але при цьому часто відбувається їх заміна (наприклад, шуба – «суба», жук – «зук»). Звук [р] найчастіше замінюється на [л], [л'] або [й] (наприклад, рука – «йука», рак – «лак», пиріжок – «пиліжок»).

- Більшості дітей без особливих труднощів удається вимова звуків [ж], [ч], [ш], [р], [р']. Але інколи трапляються випадки заміни складних звуків або неправильна їх вимова [5].

Визначимо кілька порушень вимови, які найчастіше трапляються в дітей у віці 5-6 років згідно з дослідженням М. М'ятежної: **«Ротацізм – вади вимови звуків [р] та [р'], їх відсутність або спотворена вимова. Параротацізм – різновидності дефекту, які виражаються в заміні звуків [р], [р'] іншими. Сигматизм – вади вимови звуків [с], [с'], [з], [з'], [ц], [ц'], а також вади вимови звуків [ш], [ж], [ч], [щ]. Парасигматизм – заміна звуків [ш], [ж], [ч], [щ] іншими звуками та заміна звуків [с], [с'], [з], [з'], [ц], [ц'] іншими звуками. Ламбдацізм – вади вимови звуків [л] та [л']. Параламбдацізм – різновидності дефекту, які виражаються в заміні звуків [л] та [л'] іншими»** [13].

Проаналізувавши характеристику дефектів мовлення в дикторів (див. Додаток 8), можемо зробити такі висновки:

1. Найбільш поширеними дефектами мовлення в дітей є ротацізм (параротацізм) і ламбдацізм (параламбдацізм).
2. Лише в однієї дитини з шести відсутні виражені дефекти мовлення.
3. Згідно з дослідженнями для віку мовців найбільш притаманною є складність вимовляти звуки [р] і [л].
4. Зазвичай, на одну дитину припадає 2 види дефектів мовлення.
5. Найчастіше звук [л] діти замінюють на [и], а звук [р] – на [в], [г], [й].

## ВИСНОВКИ ДО РОЗДІЛУ 2

Дитяче мовлення – це особливий феномен і важливий етап онтогенетичного розвитку мовлення людини, який надає змогу розширити наші

погляди не тільки про засвоєння мови, але також про мову в цілому, про те, як вона функціонує в мовленнєвій діяльності, про психологічні процеси, що значно розширюють наші знання про людину.

Зробивши попередній аналіз українського дитячого мовлення, ми можемо узагальнити такі особливості:

1. Загалом дитяче мовлення не має чіткого розмежування за фонетичними характеристиками на чоловіче й жіноче.
2. Форманти дитячого мовлення значно вищі, ніж у дорослому мовленні.
3. У віці 5-6 років діти найчастіше вживають однослівні синтагми.
4. П'ятьом із шести досліджених мовців притаманні дефекти мовлення.
5. Найбільш поширеними дефектами мовлення в дітей є ротацізм (параротацізм) і ламбдацізм (параламбдацізм).
6. Часто діти використовують свій власний сленг і можуть вимовляти слова у своєму особливому стилі.
7. Зазвичай дитина не може вимовити відносно довготривалий інтонаційно завершений відрізок фрази, не зробивши глибокий вдих.
8. Досить часто слова вимовляються надто швидко й нерозбірливо.
9. У дітей, зазвичай, наявні декілька інтонаційно завершених синтагм в одному реченні.
10. Монолог дітей завжди складається з довгих, розгорнутих речень, а діалог, зазвичай, характеризується короткими відповідями «так» і «ні».
11. Усе частіше в мовленні дитини наявний суржик; навіть за умови, що вона росте виключно в українськомовному середовищі. Це пояснюється впливом російськомовного контенту на телебаченні й просторах інтернету, який діти споживають щодня.
12. Уже в віці 5-6 років дитина усвідомлює загальні життєві цінності й може дати визначення поняттям «любов», «дружба», «страх».

## РОЗДІЛ 3. СТВОРЕННЯ ВЕБ-САЙТУ ДЛЯ КОРПУСУ УСНОГО ДИТЯЧОГО МОВЛЕННЯ

*У цьому розділі ми дізнаємось про процес створення веб-сайту для корпусу усного дитячого мовлення. Розглянемо його структуру, детальніше опишемо програмні засоби для його розробки, а також додаткове файлове забезпечення (аудіозаписи, файли для завантаження, анотації записів у форматі .zip, тощо).*

### 3.1. Структура веб-сайту

Веб-сайт корпусу українського усного дитячого мовлення складається з головної сторінки й чотирьох розділів (див. Додаток 1):

- Корпус.
- Файли ELAN.
- Фонетичні особливості усного дитячого мовлення.
- Розширення корпусу.

Розділи «Файли ELAN» та «Фонетичні особливості усного дитячого мовлення» містять у собі декілька підрозділів.

Далі опишемо кожен розділ більш детально:

1. **Головна сторінка сайту.** Містить основну інформацію про дослідження дитячого мовлення, мету створення корпусу, методи дослідження, реалізацію, практичне втілення, інструменти тощо. На цій сторінці користувач може ознайомитися з дослідженнями першого в Україні корпусу усного дитячого мовлення. Також на сторінці є вказівки про те, де користувач може знайти потрібні файли й інформацію.
2. **Корпус.** Сторінка розповідає про процес реалізації корпусу усного дитячого мовлення. Також тут користувач може прослухати звукові записи, із яких формувався корпус.
3. **Файли ELAN.** На сторінці в відкритому доступі містяться файли корпусу у форматі .zip, які користувач може завантажити. У ході створення корпусу ми користувалися програмою ELAN, у якій кожен звуковий запис

диктора у форматі *.wav* був посеgmentований і проанотований. Для кожного файлу ми перетворювали звуковий запис зі *Stereo* в *Mono* та створювали текстовий файл *.eaf*, щоб працювати з текстом. Аби користувач мав змогу детальніше ознайомитися з даними результатами дослідження, на цій сторінці були створені такі підрозділи: сегментація записів, позначення для анотування, особливості анотування.

4. **Фонетичні особливості усного дитячого мовлення.** На цій сторінці містяться результати досліджень фонетичних особливостей усного дитячого мовлення. Ознайомитися з ними можна в таких підрозділах:

- Особливості формант голосних звуків.
- Характеристика інтенсивності дитячого мовлення.
- Частота основного тону дитячого мовлення.
- Тривалість синтагм дитячого мовлення.
- Характеристика дефектів мовлення в дітей.
- Узагальнені фонетичні особливості українського дитячого мовлення.

5. **Розширення сайту.** Цей розділ знаходиться в процесі розробки й у майбутньому надасть можливість науковцям поповнювати корпус і додавати результати своїх досліджень усного дитячого мовлення.

Структура сайту влаштована так, щоб надати змогу науковцям вільно користуватися файлами українського корпусу дитячого мовлення й завантажувати їх, а також додавати на сайт результати своїх досліджень, щоб постійно розширювати корпус.

### **3.2. Програмні засоби для створення сайту: HTML і PYTHON**

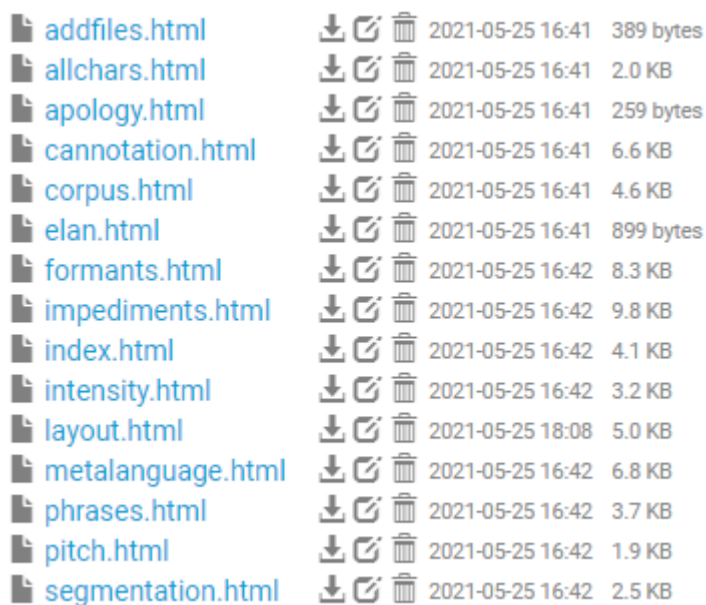
Веб-сайт був створений за допомогою двох програмних засобів:

1. HTML для *front-end*, тобто для користувацького інтерфейсу. Сторінки з розширенням *.html*, які відкриваються користувачеві в браузері, написані саме цією мовою гіпертекстової розмітки.

2. *Python* для *back-end*, тобто програмного вмісту сайту. Мовою *Python* написані функції, які обробляють інформацію надіслану сервером про дії користувача, а саме – натискання на посилання й перехід на інші сторінки.

Файли HTML-сторінок зберігаються в папці *corpussite/templates/* і завантажуються після натискання користувачем на відповідне посилання на сайті й запуску відповідної функції на програмному кінці сайту (див. Додаток 3).

Головний файл *back-end* забезпечення сайту за традицією має назву *application.py*. Він містить функції, які забезпечують правильне реагування сайту на натискання користувачем посилань на сторінках і завантаження потрібних HTML-сторінок. Розгортання серверу сайту здійснюється за допомогою популярної спеціалізованої бібліотеки *Flask*.



addfiles.html	2021-05-25 16:41	389 bytes
allchars.html	2021-05-25 16:41	2.0 KB
apology.html	2021-05-25 16:41	259 bytes
cannotation.html	2021-05-25 16:41	6.6 KB
corpus.html	2021-05-25 16:41	4.6 KB
elan.html	2021-05-25 16:41	899 bytes
formants.html	2021-05-25 16:42	8.3 KB
impediments.html	2021-05-25 16:42	9.8 KB
index.html	2021-05-25 16:42	4.1 KB
intensity.html	2021-05-25 16:42	3.2 KB
layout.html	2021-05-25 18:08	5.0 KB
metalanguage.html	2021-05-25 16:42	6.8 KB
phrases.html	2021-05-25 16:42	3.7 KB
pitch.html	2021-05-25 16:42	1.9 KB
segmentation.html	2021-05-25 16:42	2.5 KB

Малюнок 3.1. Вміст папки *corpussite/templates/*.

### 3.3. Jinja2 – система веб-шаблонів для створення .html файлів

Поряд із двома основними програмними засобами – *Python* і HTML – сайт також використовує *jinja2*.

*Jinja2* – система веб-шаблонів/шаблонізатор/рушій шаблонів для мови програмування *Python*, який дозволяє динамічно створювати HTML-сторінки.

HTML-сторінка написана лише за допомогою HTML – статична. Її вміст не може змінюватися від дій користувача, і вона може містити лише ту інформацію, яка безпосередньо записана в її *.html* файлі. *Jinja2* робить можливим керування вмістом *.html* файлів через програмний файл *.py*. Система шаблонів дозволяє передавати інформацію з файлів *.py* у вигляді змінних і розміщувати її на сторінці, а також об'єднувати різні HTML-сторінки.

При створенні сайту нам знадобився останній згаданий вище функціонал – об'єднання різних *.html* файлів.

Це пов'язано з тим, що кожна сторінка, яка існує на сайті й доступна користувачам, повинна містити в собі такі головні елементи: логотип сайту, навігаційну стрічку з розділами сайту, футер (із зазначенням автора сайту). Щоб не копіювати та вставляти один і той самий шматок HTML-коду в кожен *.html* файл, зберегли його в окремому файлі *layout.html*. Саме з цим файлом об'єднується кожна HTML-сторінка при її завантаженні в браузері користувача. Він містить зазначені вище головні елементи сторінки сайту. У кожному *.html* файлі в першому рядку вказано, що він повинен об'єднуватися з *layout.html* («розширювати» його):

```
{% extends 'layout.html' %}
```

Засобами *jinja2* у файлі *layout.html* позначені секції, де при об'єднанні буде розміщуватися HTML-код іншої сторінки.

Такими секціями є:

### **1) Заголовок сторінки** (див. Малюнок 3.2.).

Позначками `{% block title %}{% endblock %}` позначене місце для динамічної частини заголовку. Наприклад, для головної сторінки на цьому місці буде текст «Головна». При переході на сторінку корпусу текст зміниться на «Корпус». Це досягається за рахунок, по-перше, згаданих вище позначок у файлі *layout.html*, по-друге, у кожному *.html* файлі присутній такий код:

```
{% block title %}
```

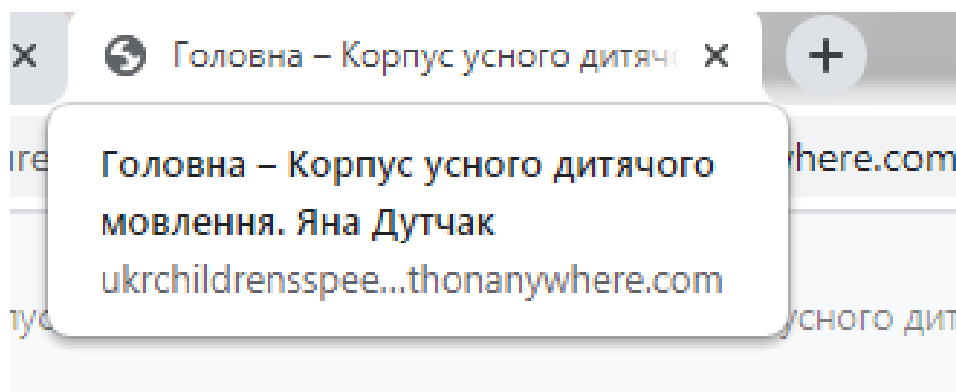
Головна  
{% endblock %}

Тут Головна – текст, який вставляється в динамічну частину заголовку тексту.

Друга частина заголовка – статична: – Корпус усного дитячого мовлення. Яна Дутчак (див. Малюнок 3.2.). Ця частина тексту буде в заголовку на кожній сторінці.

## 2) Тіло сторінки.

Принцип об'єднання коду такий самий, як і для заголовку, але використовуються позначки {% block main %}{% endblock %}.



Малюнок 3.2. Заголовок сторінки.

## 3.4. Додаткове файлове забезпечення: аудіозаписи, файли для завантаження, анотації записів у форматі .zip, тощо

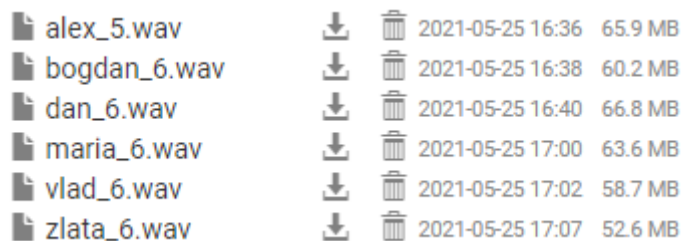
Сайт також використовує у вмісті сторінок сторонні файли: аудіозаписи корпусу, скріншоти й матеріали для завантаження (файли анотації записів у форматі .zip). Налаштування для форматування інформації на сторінках зберігаються в окремо відведеному файлі *styles.css*. Усі сторонні файли знаходяться в папці *corpussite/static/* (див. Малюнок 3.3.).



*Малюнок 3.3.* Вміст папки *corpussite/static/*.



*Малюнок 3.4.* Вміст папки *static/downloads/*.



*Малюнок 3.5.* Вміст папки *static/audio/*.

## ВИСНОВКИ ДО РОЗДІЛУ 3

У цьому розділі ми описали увесь процес створення веб-сайту для корпусу усного дитячого мовлення. Описали його структуру, вияснили, що програмні ресурси HTML та PYTHON є досить зручними для створення веб-сторінок.

Також ми описали систему веб-шаблонів *Jinja2*, яка знадобилась нам, щоб мати змогу керувати вмістом *.html* файлів.

На завершення ми окреслили додаткове файлове забезпечення, що містить сайт : аудіозаписи, файли для завантаження, анотації записів у форматі *.zip*, тощо. Таким чином, за допомогою усіх вищевказаних кроків, ми розробили веб-сайт для корпусу усного дитячого мовлення.

## **ВИСНОВОК**

На базі знань, передбачених навчальною програмою «Прикладна (комп'ютерна) лінгвістика та англійська мова», із дібраних і спеціально створених аудіозаписів за допомогою програм ELAN і PRAAT було досліджено характеристики українського дитячого мовлення й створено перший в Україні корпус усного дитячого мовлення. Спочатку було описано й досліджено теоретичні засади створення моделі усного дитячого мовлення, а згодом на невеликому за обсягом посегментованому й проанотованому матеріалі записів дитячого мовлення загальною тривалістю 64 хв було досліджено формантну структуру, частоту основного тону, інтенсивність, побудовано інтонаційні контури, описано особливості тривалості синтагм і дефектів дитячого мовлення. Далі за допомогою програмних засобів HTML і PYTHON було створено веб-сторінку для винесення результатів досліджень на широкий загал.

Підсумовуючи дані, розглянуті в даній роботі, можна зробити такі висновки:

1. Фонетичні характеристики мовлення дитини значно відрізняються від фонетичних характеристик мовлення дорослих.
2. Мова дитини розвивається не в монологі, а в діалозі – у комунікації з дорослими.
3. Створення й дослідження корпусів усного дитячого мовлення є важливим аспектом вивчення формування мовлення людини в цілому.
4. Усна форма мовлення, безумовно, є основною, фундаментальною вихідною.
5. Корпусні лінгвістичні дослідження повинні включати дослідження корпусів усного мовлення, як один із центральних компонентів.
6. Програми ELAN і PRAAT вважаються одними з найкращих за своєю базою для створення й дослідження акустичних корпусів.
7. Вік 5-6 років є найбільш оптимальним для дослідження дитячого мовлення, адже саме тоді людина вчиться контролювати власну вимову, може оволодіти монологом, здатна свідомо осмислювати певні

загальні цінності й попри типові дефекти в вимові, формує свій власний стиль мовлення.

8. За допомогою дослідження фонетичних показників дитячого мовлення можна визначити, як влаштовані дійсні мовні правила, де аномалії в мові, де закономірності, що логічно, а що – ні.
9. У кожної дитини свої унікальні фонетичні особливості, проте результати наших спостережень дають змогу скласти уявлення про певні норми.
10. Для створення веб-сайту найбільш доцільно користуватися такими програмними засобами, як *Python* та HTML.
11. Паралельно з двома основними програмними засобами – *Python* та HTML – ми використовували систему веб-шаблонів *jinja2*, яка дає змогу динамічно створювати HTML-сторінки й керувати їхнім змістом.
12. При створенні сайту нам знадобилася одна з функцій *jinja2* – об'єднання різних файлів формату *.html*.

Отже, за допомогою матеріалу, який представлений та описаний у цій роботі, можна надалі збільшувати обсяг корпусу й удосконалювати його наповнення. Також можна розширювати базу дикторів різного віку, наприклад, за місцем проживання або віком, користуючись уже наявним відкритим веб-ресурсом корпусу усного дитячого мовлення. У майбутньому ми плануємо додати на веб-сторінку функцію, яка дала б змогу всім охочим не лише користуватись її вмістом, але й вільно додавати результати своїх досліджень на сайт.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Брызгунова Е. А. Практическая фонетика и интонация русского языка. – М., 1963.
2. Вайнштейн А. М. Принципы и методика лингвистического описания детской речи / А. М. Вайнштейн // Детская речь и пути ее совершенствования: межвуз. сб. науч. трудов. Свердловск, 1989. – С. 42.
3. Выготский Л. С. Вопросы детской психологии / Л. С. Выготский. – СПб., М., 1997. – 467 с.
4. **В. Н. Ярцева. «Лингвистический энциклопедический словарь» // Научно-редакционный совет издательства «СОВЕТСКАЯ ЭНЦИКЛОПЕДИЯ». Институт языкознания АН СССР. – URL: <http://www.tapemark.narod.ru/les/> (дата звернення 07.06.2021).**
5. Вікові характеристики мовленнєвого розвитку // Рівненський навчально-виховний комплекс № 1. – URL: <http://nvk1.rosvita.rv.ua/index.php/dlia-batkiv/storinka-lohopeda> (дата звернення 07.06.2021).
6. Воротинцева О. «Основні етапи розвитку мовлення в дітей» // Батькам про дітей. Поради логопеда №3. – URL: [http://zakinppo.org.ua/images/2020/docs/05/Батькам\\_про\\_дітей\\_3.pdf](http://zakinppo.org.ua/images/2020/docs/05/Батькам_про_дітей_3.pdf) (дата звернення 07.06.2021).
7. Гридина. Т. А. Онтолингвистика. Язык в зеркале детской речи: учеб. пособие / Т. А. Гридина. – М.: Наука: Флинта, 2006. – 152 с.
8. Дубинина, И. Е., С. А. Маламуд, и Е. В. Денисова-Шмидт. «Аудио-синхронизированный многолетний корпус билингвальных Русскоязычных детей и их семей (BiRCh Longitudinal)». 2013-present. – URL: [http://people.brandeis.edu/~smalamud/MILa/birch\\_ru.html](http://people.brandeis.edu/~smalamud/MILa/birch_ru.html) (дата звернення 07.06.2021).
9. Е. В. Зырянова. Система Childe как метод сбора материалов и изучения детской речи – Пермь: 2008. – (Известия Российского государственного педагогического университета им. А .И. Герцена).

- 10.И. П. Ющук // Українська мова: Підручник для студентів філологічних спеціальностей вищих навчальних закладів/ – К.: Либідь, 2004. – 637 с.
- 11.Исенина Е. И. К вопросу о формировании образа-слова // Вопросы психологии. – 1967. – №1.
- 12.Кибрик Андрей Александрович. К созданию корпусов устной детской речи. – д.ф.н. Институт языкознания РАН Подлесская Вера Исааковна, д.ф.н. Институт лингвистики РГГУ.
- 13.Мария Мятежка. «Правильна артикуляція звуків с, з, ц. Порушення вимови свистячих звуків. Виправлення вад вимови свистячих звуків» // Артикуляція. – URL: <https://pandia.org/text/78/090/5445.php> (дата звернення 07.06.2021).
- 14.**Плахотникова Е. Ю. Особенности транскрибирования украинской устной речи в программе ELAN / Е. Ю Плахотникова // Балтийский гуманитарный журнал. 2015. No 4(13) . – с.40-43.**
- 15.Селіванова О. Сучасна лінгвістика: термінологічна енциклопедія. – Полтава Довкілля – К., 2006.
- 16.Створення акустичного корпусу українського ефірного мовлення / [Н.Б. Васильєва, В.В. Пилипенко, О.М. Радущкий та інш.]. – Оброблення сигналів і зображень та розпізнавання образів: Десята Всеукраїнська міжнародна конференція. – Київ, 2010. – С. 55-58.
- 17.Рулёва Е. А. Типология детских номинаций во французском, греческом и украинском языках / Е.А. Рулёва // Записки з романо-германської філології; [Під ред. І.М. Колегаєвої]. – Вип. 9. Одеса: Латстар, 2001. – С. 160 – 172.
- 18.Торсуева И. Г. Интонация // Лингвистический энциклопедический словарь / Главный редактор В. Н. Ярцева. – М.: Советская энциклопедия, 1990. – 685 с.
- 19.**A New Method for Collection and Annotation of Speech Data of Atypically Developing Children – URL: <https://ieeexplore.ieee.org/document/8615918> (дата звернення 25.05.2021)**

20. Bienvenue sur CLAPI – URL: <http://clapi.ish-lyon.cnrs.fr> (дата звернення 28.05.2021).
21. Boersma, Paul; van Heuven, Vincent (2001). «Speak and unSpeak with Praat» (PDF). *Glott International*. 5 (9/10): 341–347.
22. CEFC – URL: <https://www.ortolang.fr/market/corpora/cefc-orfeo> (дата звернення 28.05.2021).
23. CLARIN – European Research Infrastructure for Language Resources and Technology c/o Utrecht University/ The Netherlands – URL: <https://www.clarin.eu/resource-families/spoken-corpora> (дата звернення 07.06.2021).
24. Collen Richey / Speech Corpora: Ling 110 / 20 April 20 / Stanford University, Stanford, California 94305. – URL: [https://web.stanford.edu/dept/linguistics/corpora/material/X\\_Speech\\_Corpora.pdf](https://web.stanford.edu/dept/linguistics/corpora/material/X_Speech_Corpora.pdf) (дата звернення 07.06.2021).
- 25. Comprehending Real Numbers: Development of Bengali Real Number Speech Corpus – URL: [https://www.researchgate.net/publication/324057513\\_Comprehending\\_Real\\_Numbers\\_Development\\_of\\_Bengali\\_Real\\_Number\\_Speech\\_Corpus](https://www.researchgate.net/publication/324057513_Comprehending_Real_Numbers_Development_of_Bengali_Real_Number_Speech_Corpus) (дата звернення 28.05.2021)**
26. Dokumentenserver Hosting der SUB-Hamburg – URL: <https://edoc.sub.uni-hamburg.de/> (дата звернення 28.05.2021).
27. ELFA PROJECT – URL: <https://elfaproject.wordpress.com/tag/elfa-corpus/> (дата звернення 28.05.2021).
28. **English Broadcast News Speech Recognition by Humans and Machines – URL: <https://arxiv.org/abs/1904.13258> (дата звернення 28.05.2021).**
29. Eskenazi, Maxine, Jack Mostow, and David Graff. The CMU Kids Corpus LDC97S63. Philadelphia: Linguistic Data Consortium, 1997. – URL: <https://catalog.ldc.upenn.edu/LDC97S63> (дата звернення 07.06.2021).

30. Estonian North Wind and the Sun Corpus v.1.0.3 – URL: <https://doi.org/10.15155/1-00-0000-0000-0000-00129L> (дата звернення 28.05.2021).
31. Hamburg Modern Times Corpus (HaMoTiC) – URL: <https://corpora.uni-hamburg.de/hzsk/de/islandora/object/spoken-corpus:hamotici> (дата звернення 28.05.2021).
32. International standard language resource number – URL: <http://islrn.org/resources/999-417-152-218-3/> (дата звернення 28.05.2021).
33. ISCA – URL: [https://www.isca-speech.org/archive/Interspeech\\_2019/pdfs/2980.pdf](https://www.isca-speech.org/archive/Interspeech_2019/pdfs/2980.pdf) (дата звернення 28.05.2021).
34. KORP – URL: <http://urn.fi/urn:nbn:fi:lb-2015050502> (дата звернення 28.05.2021).
35. MacWhinney B. The CHILDES Project: Tools for Analyzing Talk. 3rd Ed. New York, 2000. – URL: <https://childes.talkbank.org/topics/> (дата звернення 07.06.2021).
36. Praat: doing phonetics by computer – URL: <https://www.fon.hum.uva.nl/praat/> (дата звернення 28.05.2021).
37. Reconceptualizing practice with multilingual children with speech sound disorders: people, practicalities and policy – URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/1460-6984.12112> (дата звернення 07.06.2021).
38. Santa Barbara Corpus of spoken American English. DuBois, John (dir.) 2000. Pt. 1-3. LDC. – URL: <https://www.linguistics.ucsb.edu/research/santa-barbara-corpus> (дата звернення 07.06.2021).
39. Speech Perception Skills of Children With Speech Sound Disorders: A Systematic Review and Meta-Analysis – URL: [https://pubs.asha.org/doi/pdf/10.1044/2019\\_JSLHR-S-18-0519](https://pubs.asha.org/doi/pdf/10.1044/2019_JSLHR-S-18-0519) (дата звернення 07.06.2021).

40. The ICSI Meeting Recorder Dialog Act (MRDA) Corpus – URL: <https://www.aclweb.org/anthology/W04-2319.pdf> (дата звернення 28.05.2021).
41. The language archive – URL: <http://www.lat-mpi.eu/tools/elan/elanforum/post201102145120546748/> (дата звернення 28.05.2021).
42. The Language Bank of Finland – URL: <http://urn.fi/urn:nbn:fi:lb-201708251> (дата звернення 28.05.2021).
43. The Machine-Readable Corpus of Spoken English – URL: <http://korpus.uib.no/icame/manuals/SEC/> (дата звернення 28.05.2021).
44. What Do Children with Speech Sound Disorders Think about Their Talking? – URL: <https://www.thieme-connect.com/products/ejournals/abstract/10.1055/s-0039-1677760> (дата звернення 07.06.2021).
45. **Word Frequencies in Written and Spoken English: based on the British National Corpus** – URL: <http://ucrel.lancs.ac.uk/bncfreq/> (дата звернення 28.05.2021).

*Додаток 1*

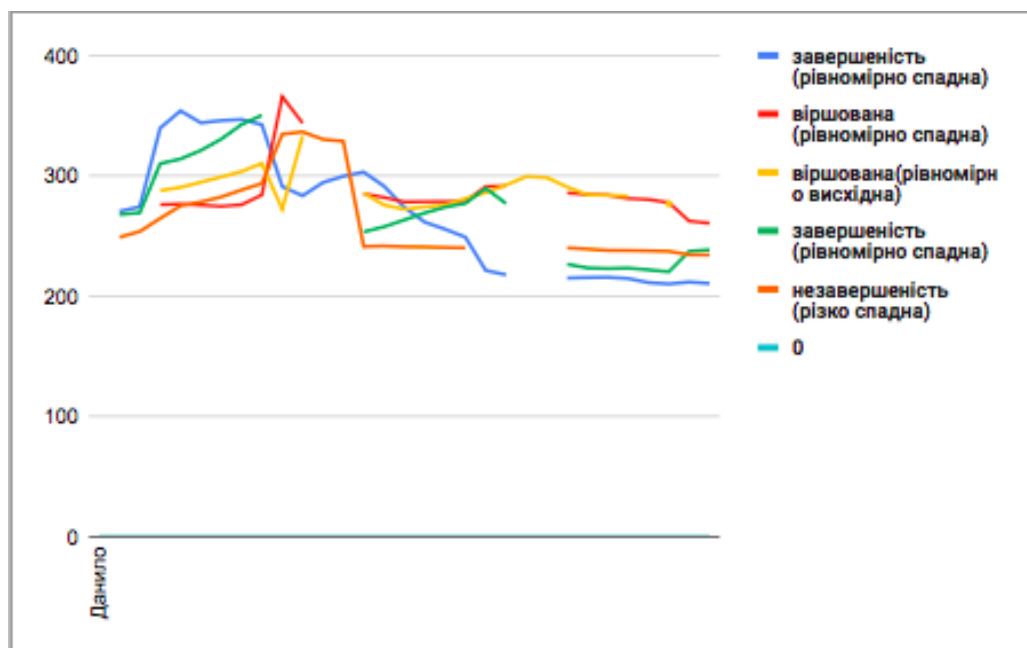
*Посилання на веб-сайт корпусу дитячого усного мовлення*

За цим посиланням знаходиться сайт українського корпусу усного дитячого мовлення, який складається з чотирьох розділів і декількох підрозділів:

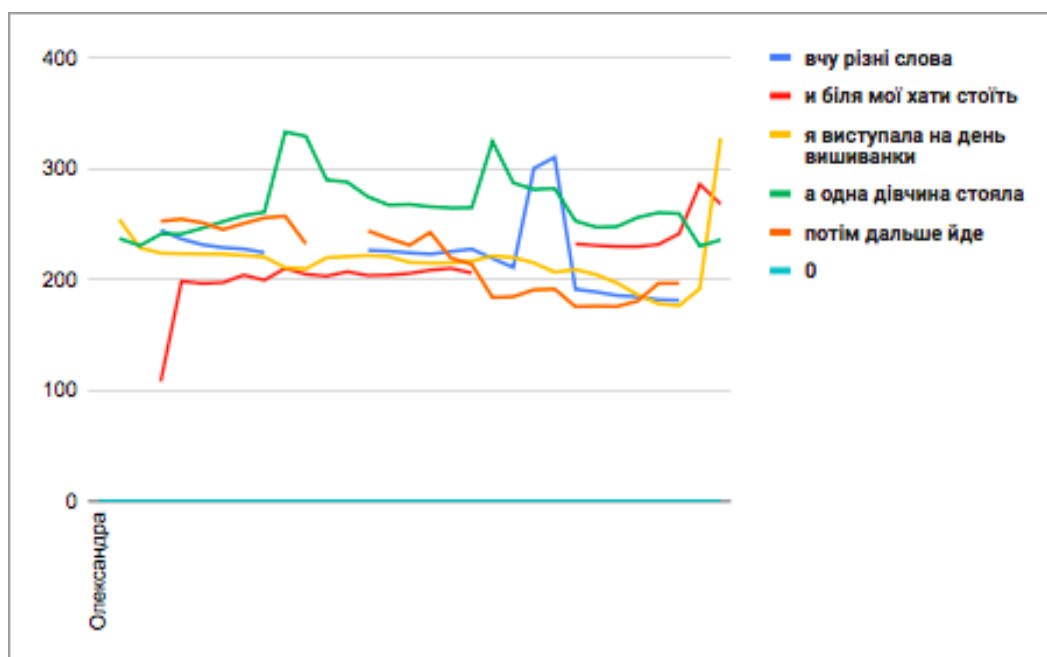
<http://ukrchildrengspeechcorpus.pythonanywhere.com>

*Графік інтонаційних контурів мовлення кожного диктора*

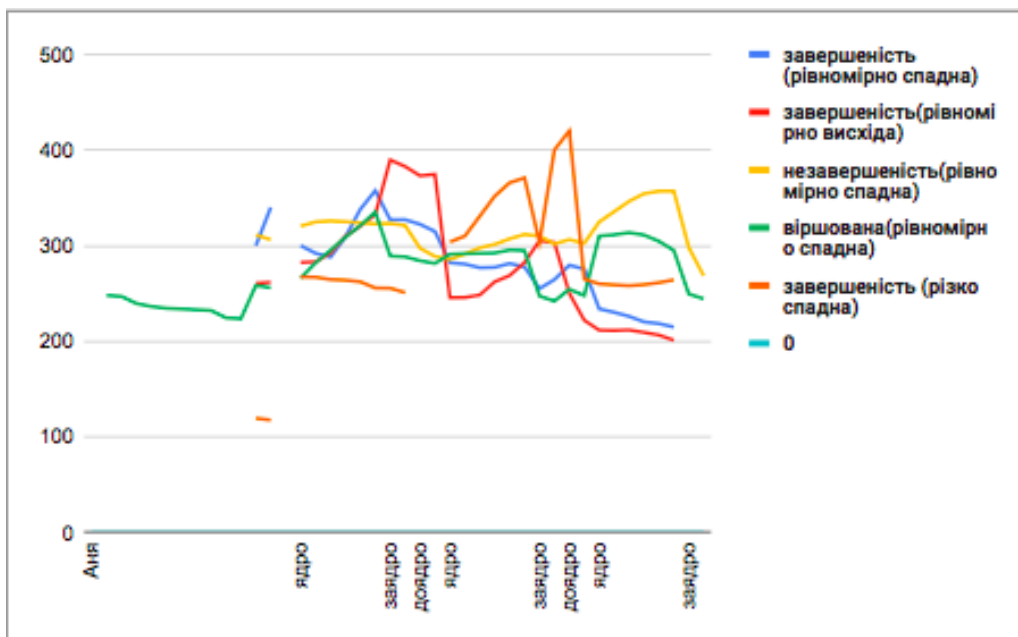
Подані графіки відтворюють особливості інтонаційних конструкцій кожного мовця.



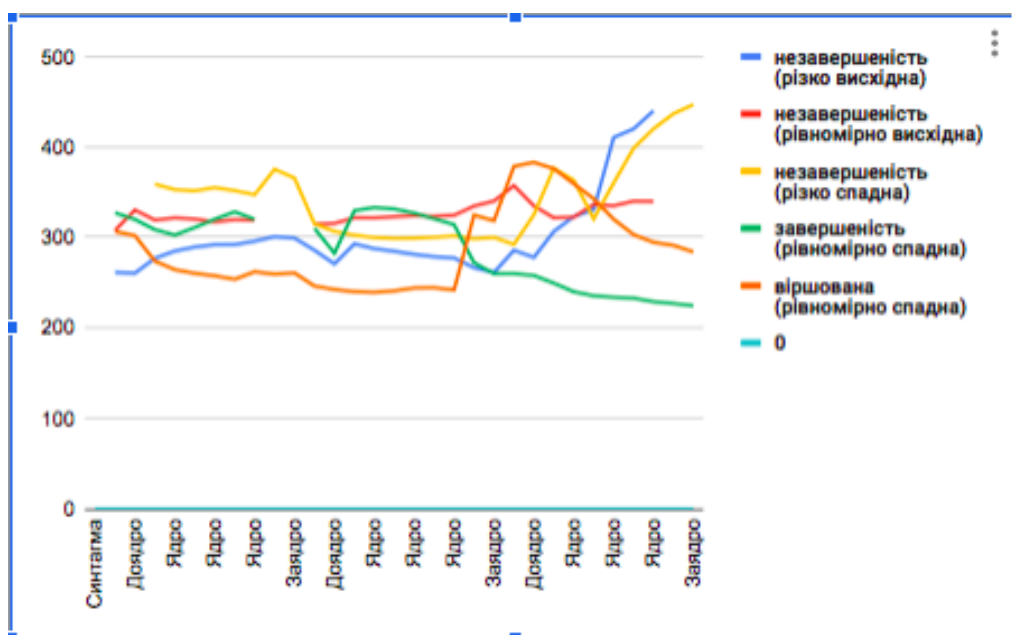
*Малюнок 1. Данило.*



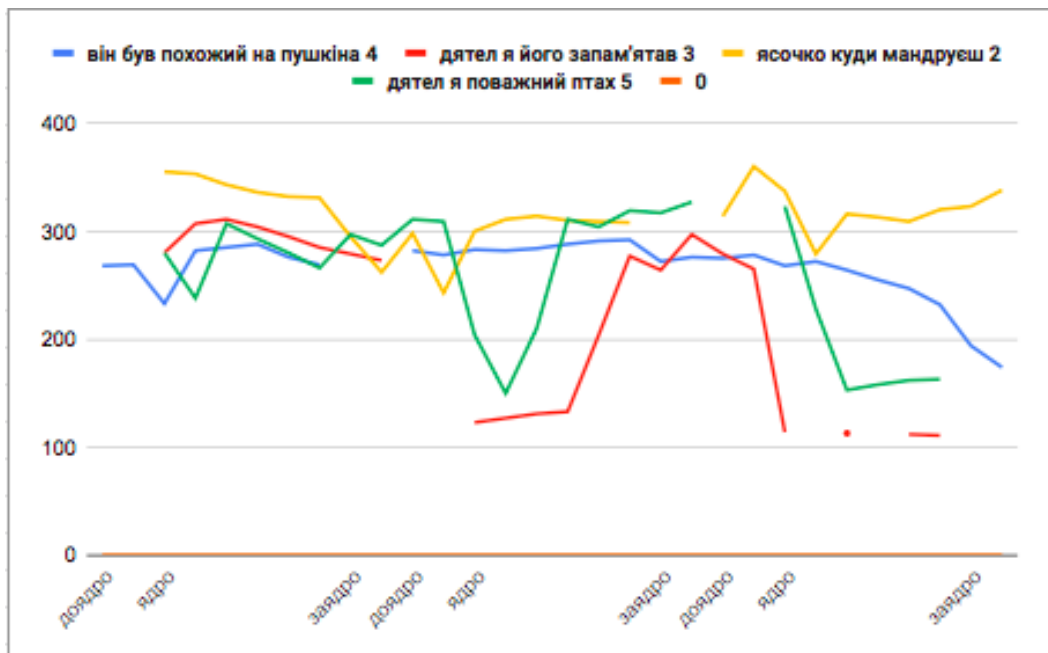
*Малюнок 2. Олександра.*



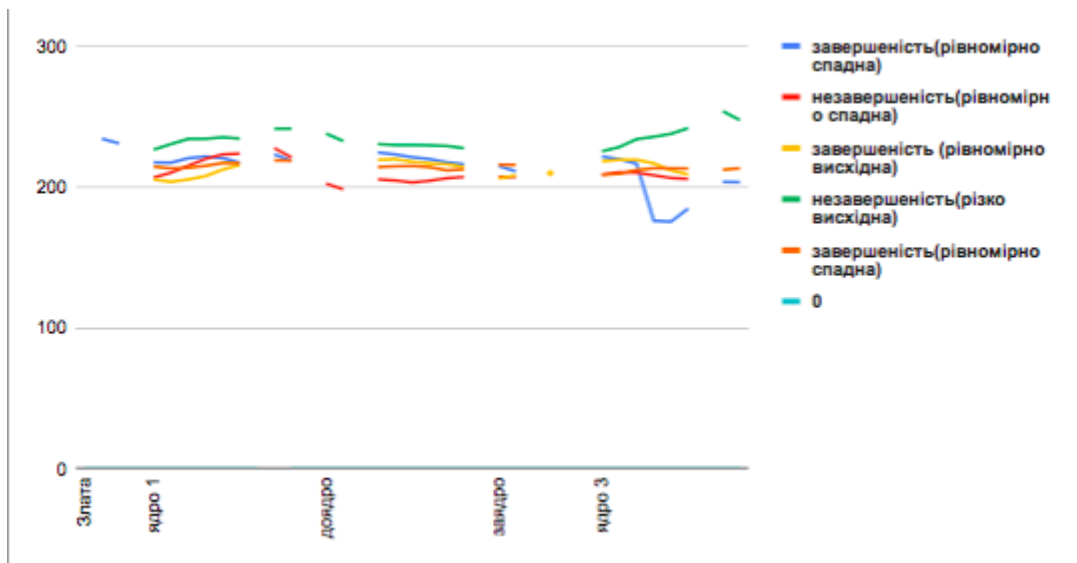
Малюнок 3. Аня.



Малюнок 4. Влад.



Малюнок 5. Богдан.



Малюнок 6. Злата.

Доступ до папки з програмними кодами знаходиться за посиланням:

<https://drive.google.com/drive/folders/1kD2lQojsMAe3nf4o7hMGMk236yxvMAvB?usp=sharing>

Якщо перейти за цим посиланням, можна побачити папку *corpussite*, у якій знаходиться 2 папки та 2 файли у форматі *.py*:

- *aplication.py* – головний файл *back-end* забезпечення сайту. Він містить функції, які забезпечують правильне реагування сайту на натискання користувачем посилань на сторінках і завантаження потрібних HTML-сторінок.
- *static* – усі сторонні файли, де *styles.css* – налаштування для форматування інформації на сторінках.
- *templates* – файли HTML-сторінок, які завантажуються після натискання користувачем на відповідне посилання на сайті й запуску відповідної функції на програмному кінці сайту.
- *helpers.py* – функція, яка опрацьовує повідомлення про помилки коду. Якщо сторінку сайту не можна відкрити через певну помилку в коді, то функція повертає повідомлення на сторінці сайту про те, що є помилка, замість того, щоб повертати незрозумілий програмний код помилки.

Для анотації звукових записів ми використовували такі позначення:

1. ставиться перед відповідним словом або якщо слів, які потребують цих позначень, декілька поруч – беремо їх у дужки:

\_д\_ – діалект;

\_с\_ – суржик;

\_о\_ – обрив: слово обривається (обрізаний звуковий файл, диктора перебивають, диктор не договорило тощо), пишемо лише ту частину слова, яка реально звучить. Якщо слово, яке обірвалося в першій синтагмі, продовжується в другій, знову ставимо «\_о\_» та дописуємо решту слова;

\_рдц\_ – редукція: слово вимовляється зі значною редукцією (пишемо літературний варіант слова);

\_роз\_ – розтягування: слово вимовляється зі значним розтягуванням одного (декількох) звуків (пишемо літературний варіант слова);

\_пскл\_ – поскладова вимова: слово вимовляється поскладово (пишемо літературний варіант слова);

\_шеп\_ – шепелявість (пишемо літературний варіант слова);

\_карт\_ – картавість (неправильна вимова звуків [р] і [л]) (пишемо літературний варіант слова);

\_об\_ – обмовка: диктор помиляється й одразу сам себе виправляє (пишемо лише ту частину слова, яка звучить в ідеалі);

\_сл\_ – сленг (може бути індивідуальна особливість вимови певного слова диктором);

\_нечіт\_ – слово вимовляється тихо, нечітко і розмито;

\_нечіт\_() – неможливо розпізнати конкретне слово диктора;

\_з\_ – заїкання;

\_шпт\_ – шепіт;

**\_см\_** – сміх (диктор вимовляє слово через сміх);

**«слово1»/«слово2»** – диктор 1 перебиває диктора 2 на певному слові;

**\_вд\_** – вдих/видих (диктор робить глибокий вдих/видих перед тим, як вимовити слово).

2. позначення стосується цілого сегмента (рядка) і ставиться на початку сегмента, фон запису диктора:

**\_кт\_** – кашель (фон);

**\_ст\_** – стукіт (фон);

**\_хор\_** – хор (декілька дикторів одночасно) (фон) – якщо ми можемо розпізнати мовлення кожного диктора, ставимо між ними «/»;

**\_см\_** – сміх (фон);

**\_вд\_** – голосний вдих/видих (фон);

**\_фшум\_** – шум на фоні розмови диктора (позначається перед початком слова з шумом і продовжується до кінця синтагми);

**\_стгн\_** – стогін;

**\_см\_** – сміх;

**\_нечіт\_** – неможливо розпізнати жодного слова диктора;

**«текст1»/«текст2»** – диктори говорять одночасно, текст стенограми дикторів записуємо через «/»;

**роз(текст)** – розтягується весь сегмент;

**\_інт\_** – диктор говорить з однаковою інтонацією (вірш).

3. позначення означає відповідні звуки у записі, які вимовляє диктор (ставиться на місці відповідного звука):

**\_см\_** – сміх;

**\_к\_** – кашель;

**\_вд\_** – вдих/видих;

**\_а\_** – а-кання;  
**\_е\_** – е-кання;  
**\_м\_** - ме-кання;  
**\_і\_** – і-кання;  
**\_и\_** – и-кання;  
**\_схл\_** – схлипування;  
**\_угу\_** угу-кання;  
**\_гм\_** гм-кання;  
**\_скр\_** – скрип;  
**\_цьм\_** – цьмя-кання;  
**\_ем\_** довге «-ем»;  
**\_запереч\_у** – заперечний звук «-у-у».

4. позначення стосується цілого сегмента (позначається конкретний вид шуму без голосу диктора):

**\_вд\_** – вдих/видих;  
**\_цьм\_** – цьмя-кання;  
**\_каш\_** – кашель;  
**\_стук\_** – стукіт;  
**\_шум\_** – шум невизначеного походження;  
**#** – тиша, пауза.

***Додаток 5***

***Дані таблиць щодо середнього значення форм першої та другої форманти голосних дитячого мовлення***

Подані таблиці є результатом аналізу декількох наголошених голосних у кожного диктора та їх середнього значення. Ми занесли результати в окрему таблицю для кожного диктора.

Голосний звук	а	о	у	е	и	і
Перша форманта	943.8	704.0	492.2	823,3	762.3	490.5
Друга форманта	1252.8	1389.2	1363.7	1574.2	1801.7	1786

**Таблиця 1.**

**ДИКТОР: АНЯ**

Голосний звук	а	о	у	е	и	і
Перша форманта	1002.6	661.2	501.9	927.9	633.8	574.7
Друга форманта	1676.7	1145.2	1397.5	2216.5	2528.1	2538.7

**Таблиця 2.**

**ДИКТОР: БОГДАН**

Голосний звук	а	о	у	е	и	і
Перша форманта	780.7	627.2	626.3	697.6	631.2	477.3
Друга форманта	1532.8	1354	1186.7	1545.2	2015.9	2157.3

**Таблиця 3.**

**ДИКТОР: ЗЛАТА**

Голосний звук	а	о	у	е	и	і
Перша форманта	884.1	734.5	558.6	795.8	674.1	424.2
Друга форманта	1244.2	1244.2	947.6	1807.7	1538.2	2268.9

**Таблиця 4.**

**ДИКТОР: ДАНИЛО**

Голосний звук	а	о	у	е	и	і
Перша форманта	757.4	594.7	602.4	672.7	657.3	518.9
Друга форманта	1602.5	1410.1	1273.8	2285.3	2192.7	1329.8

**Таблиця 5.****ДИКТОР: ОЛЕКСАНДРА**

Голосний звук	а	о	у	е	и	і
Перша форманта	635.9	682.5	469.2	814.3	436.8	567.4
Друга форманта	1645.4	1260.2	1454.5	2217.3	1356.7	2360.2

**Таблиця 6.****ДИКТОР: ВЛАД**

Голосний звук	а	о	у	е	и	і
Перша форманта	834	667,3	541,7	788,6	632,5	508,8
Друга форманта	1492,4	1300,4	1270	1941	1905,5	2073

**Таблиця 7.****СЕРЕДНЄ ЗНАЧЕННЯ ВСІХ ФОРМАНТ**

Подана таблиця є результатом обчислень мінімального, максимального й середнього значення інтенсивності мовлення за формулою.

Диктор:	min (дБ)	max (дБ)	Середнє значення (дБ)
Аня	40	82	68
Богдан	28	84	71
Злата	34	80	65
Данило	37	83	71
Олександра	34	87	70
Влад	32	84	71

**Таблиця 8.**

**ХАРАКТЕРИСТИКА ІНТЕНСИВНОСТІ ДИКТОРІВ**

Подана таблиця є результатом аналізу частоти основного тону кожного диктора в мінімальному, максимальному й середньому показниках.

Диктор:	min (Hz)	max (Hz)	Середнє значення (Hz)
Аня	132	470	306
Богдан	180	457	302
Злата	167	460	272
Данило	204	478	269
Олександра	170	480	268
Влад	136	465	302

**Таблиця 9.**

**ЧОТ КОЖНОГО ДИКТОРА**

Подана таблиця ілюструє найпоширеніші дефекти мовлення в кожного диктора.

Диктор	Дефекти мовлення	Приклади
Аня	Відсутність виражених дефектів мовлення	Відсутні
Богдан	<ol style="list-style-type: none"> <li>1. Сигматизм.</li> <li>2. Ламбдацизм (пом'якшення вимови звука [л]):</li> <li>3. Інколи ротацизм, але лише у словах, які починаються на звук [р] або стоять після глухого приголосного.</li> <li>4. Парасигматизм.</li> </ol>	<ol style="list-style-type: none"> <li>1. - всі ве[ш]няні ні, ми пишем там - в[ш]які по рисочкам - пишемо в[ш]ьо</li> <li>2. - до шко[л']и - весе[л']ці - ми співа[л']и пісеньку</li> <li>3. - в школу [р]адості, - іг[р]и г[р]аєм. - а він нам [р]оздав цукерки У слові «друзів» не спостерігається.</li> <li>4. - «[ж]ловив» замість «зловив» - ро[ж]мальовуємо - [ж]ошити і ми на них пишемо всьо</li> </ol>
Злата	<ol style="list-style-type: none"> <li>1. Параротацизм (заміна звуком [й]).</li> <li>2. Параламбдацизм (заміна звуком [и]).</li> <li>3. Ротацизм.</li> </ol>	<ol style="list-style-type: none"> <li>1. - т[й]охи хочу т[й]охи не хочу - п[й]осто я боюсь темноти - до нового [й]юкзака чи</li> <li>2. - [и]еся українка - це, я малюва[и]а - іноп[и]анетянського корабля</li> <li>3. - ко[р]абель інопланетянів) - фея п[р]илетіла забрала зуб - г[р]ивень</li> </ol>

Данило	1. Параротацізм (заміна звуком [г]). 2. Ламбдацізм. 3. Ротацізм.	1. - я їхав на чо[г]не море - м мака[г]они - і ка[г]топля 2. - п[л]авав - по [л]авці - ячка шоко[л]адні 3. - танцюємо [р]озказуємо віршики - і[р]ина м[р]итвіна - [р]озказувати ві[р]шик)
Олександр ра	1. Ламбдацізм. 2. Ротацізм.	1. - бо ко[л]и я виступа[л]а бачи[л]а - а я зошит забу[л]а вдома - фарбами деко[л]и 2. - четве[р]о дівчат - ми пишем то що там по к[р]апочках - т[р]єба обводити
Влад	1. Параротацізм: а) заміна звуком [в], коли звук стоїть на початку слова чи після глухого приголосного; б) заміна звуком [и] чи ковтання звука, коли звук стоїть у середині слова. 2. Ламбдацізм.	1. а) - він там шукав [в]озуму, - до со[в]озуму, до со[в]оки і він - його вчила к[в]асти б) - що в неї багато ку[и]чаток - _роз_до_карт_во[р]они - і го[р]обчик спитав 2. - і вона сказа[л]а - потім він по[л]етів - і вона сказа[л]а

Таблиця 10.

## ДЕФЕКТИ МОВЛЕННЯ КОЖНОГО ДИКТОРА