

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

**Аналіз просодичних характеристик мовлення білінгвів**  
**(експериментально-фонетичне дослідження)**

**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»

**Наталії МЕЛЬНИЧУК**

**Науковий керівник:**

Валентина РОБЕЙКО

**Рецензент:**

канд. філол. наук Зоя ДУДНИК

**Київ – 2021**

## ЗМІСТ

<b>ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ</b>	4
<b>ВСТУП</b>	5
<b>РОЗДІЛ 1. БІЛІНГВІЗМ ЯК МІЖДИСЦИПЛІНАРНЕ ЯВИЩЕ: ПРОБЛЕМИ ДЕФІНІЦІЇ ТА АСПЕКТИ ДОСЛІДЖЕННЯ</b>	9
1.1. Проблема дефініції та аспекти вивчення терміну «білінгвізм»	9
1.2. Білінгвізм як соціальне та психолінгвістичне явище: огляд сучасної наукової літератури	12
<b>Висновки до розділу 1</b>	15
<b>РОЗДІЛ 2. ЛІНГВІСТИЧНА ІНТЕРФЕРЕНЦІЯ В УМОВАХ БІЛІНГВІЗМУ</b>	17
2.1. Загальне поняття інтерференції, її причини та прояви на різних мовних рівнях	17
2.2. Особливості українсько-англійської просодичної інтерференції	19
<b>Висновки до розділу 2</b>	23
<b>РОЗДІЛ 3. СТВОРЕННЯ ТА АВТОМАТИЧНИЙ АНАЛІЗ КОРПУСУ МОВЛЕННЯ БІЛІНГВІВ</b>	24
3.1. Створення корпусу мовлення білінгвів	24
3.2. Синтагматичне членування записів мовлення	26
3.3. Створення скриптів для аналізу просодики	27
3.3.1. Скрипт boundaries_lv3-4.praat	28
3.3.2. Скрипти meanpitch.praat та meanintensity.praat	29
3.3.3. Скрипт mean_all.praat	30
3.3.4. Скрипт count_pauses.praat	31
3.3.5. Скрипт get_phrases.praat	32
3.4. Розробка програмних кодів для аналізу синтагм	33
3.4.1. Програма edit_phrases.py	33
3.4.2. Програма count_phrases.py	34

<b>Висновки до розділу 3</b>	35
<b>РОЗДІЛ 4. ПРОСОДИЧНІ ХАРАКТЕРИСТИКИ МОВЛЕННЯ</b>	
<b>БІЛІНГВІВ</b>	37
4.1. Особливості паузації	37
4.2. Довжина синтагм	41
4.3. Частота основного тону (ЧОТ) та інтенсивність (ІЗС)	45
4.4. Темп мовлення	47
<b>Висновки до розділу 4</b>	49
<b>ВИСНОВКИ</b>	51
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</b>	53
<b>ДОДАТКИ</b>	57
Додаток 1. Відомості про дикторів	57
Додаток 2. Список тем для спонтанного мовлення дикторів	61
Додаток 3. Скрипт boundaries_lv3-4.praat	62
Додаток 4. Скрипт meanpitch.praat	63
Додаток 5. Скрипт meanintensity.praat	64
Додаток 6. Скрипт mean_all.praat	65
Додаток 7. Скрипт count_pauses.praat	66
Додаток 8. Скрипт get_phrases.praat	67
Додаток 9. Програма edit_phrases.py	68
Додаток 10. Програма count_phrases.py	69
Додаток 11. Кількість синтагм за довжиною	70

## **ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ**

ІЗС – інтенсивність звукового сигналу

СМ – спонтанне мовлення

ЧМ – читане мовлення

ЧОТ – частота основного тону

## ВСТУП

На сучасному етапі розвитку лінгвістики та психолінгвістики особливий інтерес серед науковців викликає дослідження особливостей усного мовлення, яке вважається первинним щодо писемного й зумовлює виникнення більшості мовних змін. Причиною такого зацікавлення є бурхливий розвиток інформаційних технологій, які уможливають дослідження фонетичних і просодичних характеристик усної комунікації на сегментному та супрасегментному рівнях. Залучення спонтанного та підготовленого мовлення білінгвів до об'єктів дослідження експериментальної фонетики є закономірним та важливим завданням ще й тому, що сприяє ґрунтовному аналізу явища білінгвізму та глибшому осмисленню суті категорії двомовності.

Аналіз наукової літератури, присвяченої феномену білінгвізму, свідчить про недостатній ступінь його дослідження в лінгвістичному (акустико-фонетичному) та психолінгвістичному аспекті. Оцінюючи проблему вивчення просодії мовлення білінгвів, варто підкреслити, що окремі питання специфіки фонетичних характеристик мовлення (зазвичай артикуляційних) розглядалися на матеріалі різних мов у ґрунтовних працях вітчизняних (А.Калита, О.Валігура, Л.Тараненко та ін.) і зарубіжних вчених (У.Вайнрайх, Д.Москалева, Ф.Грожан, Е.Хауген та ін.).

**Актуальність** роботи як у теоретичному, так і в практичному плані полягає у відсутності спеціальних досліджень (а також акустичного корпусу для їх проведення) просодичних характеристик мовлення українських білінгвів і потребою його комплексного порівняльно-фонетичного аналізу. Вибір теми дослідження узагальнюється тенденціями сучасного мовознавства до глибокого, докладного вивчення просодичної організації спонтанного та читаного мовлення.

**Мета** дослідження – укласти корпус усного мовлення білінгвів та шляхом проведення аудитивного та акустико-фонетичного аналізу виявити основні

просодичні, зокрема диференційні та інтегральні ознаки українського та англійського мовлення дикторів-білінгвів.

Реалізація поставленої в роботі мети передбачає розв'язання **низки завдань**:

1. Розглянути комплекс теоретичних питань, пов'язаних із дослідженням білінгвізму у вітчизняній та зарубіжній лінгвістиці, зокрема психолінгвістиці.
2. Дослідити явище лінгвістичної, у тому числі просодичної інтерференції в умовах білінгвізму.
3. Створити акустичний корпус спонтанного та читаного мовлення українських білінгвів.
4. Здійснити аналіз просодичних параметрів українського та англійського мовлення за допомогою інструментального аналізу та спеціально розроблених скриптів.
5. Визначити варіантні та інваріантні ознаки просодичного оформлення висловлювань в українському та англійському мовленні.

**Об'єктом** дослідження є спонтанне та підготовлене українське й англійське мовлення білінгвів-українців.

**Предмет** дослідження – акустичні параметри англійського мовлення українських білінгвів у порівнянні з їхнім українським мовленням.

**Матеріалом** дослідження слугували аудіозаписи спонтанного та підготовленого мовлення дев'ятьох українських білінгвів; шість аудіозаписів були створені в лабораторії експериментальної фонетики Інституту філології Київського національного університету імені Тараса Шевченка у 2020 році із залученням трьох дикторів, а решта (22 аудіозаписи) записані в домашніх умовах за допомогою диктофонів на мобільних пристроях. Для читаного мовлення кожного диктора були підготовлені тексти із сучасної художньої прози, а саме уривки із романів “Гаррі Поттер і філософський камінь” та “Гаррі Поттер і таємна кімната” англійської письменниці Джоан Роулінг, відібрані з мережі Інтернет.

Загальний обсяг матеріалу експериментальної частини – акустичного корпусу з анотацією – складає 2 год 10 хв 25 с; тривалість аудіозаписів спонтанного мовлення становить 1 год 13 хв 24 с, читаного – 57 хв 1 с.

Мета й завдання дослідження вимагають комплексного застосування низки таких **методів**: теоретичних загальнонаукових (аналіз і синтез, індукція та дедукція, формалізація, узагальнення); експериментально-фонетичного (аудитивний та акустичний аналіз); методу емпіричних досліджень (спостереження, вимірювання, порівняння, математично-статистичний метод опрацювання числових значень за допомогою самостійно написаних скриптів у програмному середовищі Praat). Інструментальний аналіз просодичних характеристик мовлення здійснено за допомогою комп'ютерної програми Praat.

**Методологічну основу** роботи складають праці вітчизняних та зарубіжних вчених-лінгвістів в області загальної теорії мови й мовлення та білінгвізму.

**Теоретичне значення** роботи полягає в доповненні наукових знань щодо просодичної системи української та англійської мови, зокрема про закономірності взаємодії просодичних підсистем у процесі міжмовної інтерференції. Було проаналізовано та виконано системний опис основних просодичних характеристик мовлення білінгвів, виявлено інтегральні та диференційні ознаки українського та англійського мовлення.

**Практичне значення** результатів роботи полягає в можливості їх використання в подальших фонетичних дослідженнях, спеціальних курсах і семінарах з експериментальної фонетики, для поглиблення знань про акустичні особливості мовлення білінгвів. Створений акустичний корпус буде корисним у науковій роботі в галузі фонетики, корпусної лінгвістики, психолінгвістики та інших лінгвістичних дисциплін.

**Структура роботи** обумовлена визначеними метою, завданнями, об'єктом і предметом дослідження. Робота складається зі вступу, чотирьох розділів: теоретичної частини (Розділ 1. «Білінгвізм як міждисциплінарне явище: проблема дефініції та аспекти вивчення»; Розділ 2. «Лінгвістична

інтерференція в умовах білінгвізму») та практичної частини (Розділ 3. «Створення та автоматичний аналіз корпусу мовлення білінгвів»; Розділ 4. «Просодичні характеристики спонтанного та підготовленого мовлення білінгвів»), висновків, списку використаних, додатків. Загальний обсяг бакалаврської роботи становить 70 сторінок. Із них основного тексту 53 сторінки, список використаних джерел (47 найменувань) – три сторінки та додатки на 14 сторінках.

## **РОЗДІЛ 1. БІЛІНГВІЗМ ЯК МІЖДИСЦИПЛІНАРНЕ ЯВИЩЕ: ПРОБЛЕМИ ДЕФІНІЦІЇ ТА АСПЕКТИ ДОСЛІДЖЕННЯ**

### **1.1. Проблема дефініції та аспекти вивчення терміну «білінгвізм»**

Білінгвізм не є новим мовним феноменом, оскільки на даний час більшість світового населення розмовляє двома й більше мовами. Незважаючи на це, визначення поняття «двомовність» не має консенсусу, тому різні теоретичні дисципліни та міждисциплінарні напрями розробили безліч його описів і тлумачень протягом останнього століття.

Існує багато різних способів описати білінгвізм, і дослідники, як правило, приймають рішення щодо власного визначення відповідно до класифікації суб'єктів та власних цілей дослідження [22, с.3]. І досі залишається актуальною думка Р.Фрумкіної, яка зауважувала, що «термін двомовність завжди був нечітким, а в даний час його значення не тільки розмилось, а й розглядається в ореолі соціальних проблем» [15, с.165]. Погляди Фрумкіної підтримував і П.Ауер, стверджуючи, що «білінгвізм – дуже складне явище» [19, с.2], й у зв'язку з цим необхідно розрізняти сутність двомовності у соціологічному та лінгвістичному аспектах.

Традиційно виділяють чотири основних підходи до опису значення білінгвізму. На думку одних вчених [14; 2], неправомірно ставити питання про різні аспекти дослідження білінгвізму, оскільки двомовність – чисто психологічна проблема. На іншому полюсі знаходяться дослідники, які вважають, що цю проблему треба вважати власне лінгвістичною [9]. Треті у визначенні двомовності виходять суто з громадської (соціологічної) функції мови [16; 4]. Четверті пропонують лінгвокультурологічний аспект дослідження контактуючих мов [7; 11]. Крім того, зараз розвиваються й інші міждисциплінарні аспекти дослідження явища «двомовності», а саме психолінгвістичний, нейролінгвістичний, лінгвокогнітивний і т.д.

Наприкінці ХХ століття багато мовознавців почали спрямовувати свої зусилля на дослідження нової сфери, яка пізніше стане відомою як білінгвізм.

Були запропоновані різні визначення, які охоплюють весь «двомовний континуум». З одного боку знаходиться максималістський високий стандарт Блумфільда, який визначає двомовність як «володіння двома мовами на рівні носія мови» [21, с.56]. Однак пізніше дослідники спростували це твердження, вважаючи його надто категоричним; Дібольд (1964) та Макнамара (1969) запропонували більш мінімалістичний підхід, за яким білінгвомом вважається людина, яка знає чи розпізнає кілька фраз іншою мовою, тобто має тільки базові знання [25; 37]. Зайнявши більш помірковану позицію, Хауген (1953) визначив білінгвізм як «ситуацію, коли мовець може висловити цілісні змістовні думки іншою мовою» [30, с.7], і хоча це визначення є більш вичерпним, ніж визначення Дібольда та Макнамари, однак значно нижчим за стандарт Блумфільда.

«Почергове використання двох і більше мов однією і тією ж особою» було визначенням, представленим Маккі у 1962 р. [36, с.52]. Вайнрайх також погодився з Маккі, запропонувавши у 1968 р. одне з найвідоміших тлумачень: «практика поперемінного користування двома мовами буде називатися [тут] двомовністю, а особи, що її здійснюють – двомовними» [47, с.1]. Ці дослідники зробили акцент не на рівні володіння мовами та умовах їх вживання, а на використанні мов, і вже у 2000 р. цей підхід продовжили розвивати Хамерс і Бланк, які вважали, що мовець може користуватися різними мовами не однаково, проте здатність обмінюватися думками та обирати мову спілкування відповідно до часу, місця, мети та соціального контексту робить людину двомовною [29]. Пізніше Грожан (2008) у своєму визначенні також зосередив увагу на використанні мов, стверджуючи, що «білінгв використовує дві мови – окремо або почергово – для різних цілей, у різних сферах життя, з різними людьми» [28, с.14].

Майерс-Скоттон (2009) при визначенні терміну білінгвізм робить акцент на майстерності мовленнєвих навичок, вважаючи їх найважливішими: «білінгвізм – це здатність використовувати дві або більше мов на рівні, достатньому для проведення невимушеної розмови» [40, с.44].

Мабуть, найбільш ліберальне визначення було запропоновано Лі (2008), коли він визначив білінгва як «людину, яка може спілкуватися кількома мовами активним (за допомогою розмови та письма) чи пасивним (за допомогою слухання та читання) способом» [35, с.4].

Незважаючи на всі спроби вчених та інших людей визначити тлумачення білінгвізму, цей термін все одно залишається неоднозначним. У більшості випадків вищезгадані дефініції не вказують на рівень володіння мовами, тому припускається, що практика почергового використання мов вже передбачає можливість їхнього застосування у комунікативних цілях. Проте часто таке визначення є недостатнім, так як рівень володіння мовами, для прикладу, є головним критерієм з лінгвістичної точки зору [12].

Незважаючи на те, що явище «білінгвізму» охоплює стільки дисциплін та дослідницьких теорій, спроби окреслити універсальне визначення двомовності – це як відкрити скриньку Пандори [33]. Тому білінгвізм продовжує бути феноменом без реального консенсусу щодо свого тлумачення. Американський нейропсихолог А.Арділа назвав білінгвізм «занедбаною та хаотичною областю», тому досить сумнівним є те, що буде розроблене єдина чітка дефініція для даного терміна [18, с.31].

Мацнева Є.А. у своїй роботі стверджує, що «попри наявність великої кількості визначень, очевидним є те, що для усіх дефініцій терміна «білінгвізм/двомовність» спільною є сема «співіснування двох мов», а для дефініцій терміна «білінгв» – сема «двомовний індивід, що володіє двома мовами», а нанизування додаткових сегментів визначається підходом до вивчення питання» [12, с. 40].

Беручи до уваги всі проаналізовані праці та дефініції терміну «білінгвізм/двомовність», у нашому дослідженні білінгва визначаємо як людину, яка почергово використовує дві і більше мови на рівні, достатньому для комунікації та висловлення власних думок.

## 1.2. Білінгвізм як соціальне та психолінгвістичне явище: огляд сучасної наукової літератури

Як уже зазначалося, визначення терміну «білінгвізм» мало певний поступовий розвиток. Початкові концепції представляються радикальними постулатами, які підкреслюють ідею «збалансованої» людини-білінгва, що повністю володіє обома мовами. З іншого боку, сучасні дефініції перестають класифікувати чи вішати певні ярлики на термін «білінгв», а скоріше зосереджуються на описі або конкретизації психологічного розвитку, когнітивної організації чи впливу культурного середовища на двомовну особистість.

З соціокультурної точки зору виділяють чотири типи білінгвізму: стабільний, динамічний, перехідний та рудиментарний. Стабільний білінгвізм стосується лінгвістичної диференціації, яка відбувається між двома групами, що поділяють спільну територію, і тому білінгвальна група зобов'язана розрізняти використання однієї мови та іншої відповідно до соціолінгвістичної сфери. Динамічний білінгвізм розмежовує ролі та використання різних мов у соціальній ситуації при мовній асиміляції. Що стосується перехідного білінгвізму, він відображає перехід до виняткового використання однієї з мов відповідно до різних комунікативних функцій. Рудиментарний білінгвізм виконує символічну функцію, пов'язану з невеликою меншиною, яка незабаром вимре [26, с.224].

У даний час існує кілька теоретичних перспектив та напрямів вивчення білінгвізму. Г.Мартінес [38, с.39] розрізняє такі три основні аспекти (перспективи) дослідження двомовності: психолінгвістичний, соціолінгвістичний та критичний лінгвістичний. З лінгвістичної точки зору, об'єктом вивчення є мовні системи та їх взаємодія при білінгвізмі. У той же час критична лінгвістична перспектива не розглядає білінгвізм як здатність говорити двома мовами, але усвідомлювати соціокультурний, політичний та ідеологічний контексти, де має місце мова та її носії. Соціолінгвістичний аспект зосереджується на тому, як мовна система використовується в різних

соціальних сферах, залежно від комунікативних обставин та цілей. Нарешті, психолінгвістика білінгвізму націлена на вивчення процесів, пов'язаних із відтворенням, сприйняттям, розумінням і запам'ятовуванням мов, коли використовується білінгвальна мовна модель [13].

Психолінгвістів насамперед цікавить інтенсивність вияву білінгвізму, тобто його якісна сторона, оскільки основна увага приділяється процесам сприйняття та розуміння мов, характеру їх взаємодії в мисленні і мовленнєвій діяльності білінгва. На тему автономії першої та другої мови в свідомості білінгва проведено багато досліджень, наприклад, щодо розрізнення функціонального, комплексного та підпорядкованого білінгвізму. Згідно з цією класифікацією, існує три типи білінгвізму: функціональний (білінгв має два набори одиниць значення та два способи вираження; тобто лексеми двох мов є окремими сутностями й кожна належить до своєї мови), комплексний (білінгв має один набір одиниць значення та два способи вираження; тобто для одного і того ж значення існують «еквівалентні» лексеми в різних мовах) та підпорядкований (білінгв має одиниці значення першої мови та два способи вираження: першої мови та другої, вивченої за допомогою першої; таким чином, людина під час спілкування використовує переклад лексем з рідної мови для інтерпретації слів нерідної); останній тип білінгвізму зазвичай є результатом навчання у школі чи вищому навчальному закладі [39]. Проте, не існує переконливих доказів, що підтверджують таку класифікацію.

Багато психолінгвістичних досліджень свідчать про те, що білінгвізм має значні переваги при мовному та когнітивному розвитку дітей. Б.Бейн [20, с.8] та Дж.Каммінз [23, с.42] стверджують, що лінгвістичні та когнітивні вміння дітей, які є білінгвами, значно перевищують навички дітей, які розмовляють однією мовою. Контрастивні погляди на білінгвізм та його можливі негативні наслідки, такі як соціальна маргіналізація, інтелектуальні та емоційні труднощі, психологічні та мовні проблеми, серед іншого, втрачають свою вагомість та інтерес у галузі психолінгвістики.

Лінгвістика сама по собі має справу з білінгвізмом, пов'язаним із лінгвістичною інтерференцією та мовними контактами. Мовний контакт може мати як «масовий» характер із залученням цілих мовних спільнот (як, наприклад, в українсько-російському білінгвізмі), так і «індивідуальний» чи «груповий» (наприклад, українсько-англійський білінгвізм індивіда чи певної групи людей). Також говоримо про паралельний білінгвізм у ситуації, коли білінгви вивчають дві мови одночасно, швидше за все, проживаючи у двомовній сім'ї в дитинстві. Коли ж людина вивчає другу мову пізніше, у результаті переїзду в чужу країну чи під час навчання в школі, вживаємо термін послідовний білінгвізм [44, с.374].

Психологія ж займається дослідженням білінгвізму з точки зору засвоєння дітьми другої мови та з метою встановлення відповідного віку для початку її вивчення. Більшість білінгвальних дітей, які вивчають дві мови (одна з яких є їхньою рідною) до 5 років, не мають акценту ні в одній з мов, тоді як білінгви, що вивчають другу мову в дорослому віці, часто мають іншомовний акцент [44, с.374]. Таким чином, чим раніше людина почне вивчати другу мову, тим більша ймовірність, що вона звучатиме як її носій. У цьому плані психологія та лінгвістика поділяють спільну точку зору.

Соціолінгвістика білінгвізму аналізує різні варіації між соціальною та мовною структурою. На думку С.Ромейна [43, с.258], двомовність не можна пояснити лише зі сторони лінгвістики, її слід досліджувати набагато глибше, у межах когнітивного, соціального та культурного вимірів.

У цьому сенсі соціолінгвістична перспектива білінгвізму включає такі аспекти, як бікультуралізм, монокультуралізм, акультурація, мультикультуралізм, мовний вибір, взаємозв'язок між мовою та ідентичністю, володіння мовою та статус мови [26, с.224].

При аналізі явища білінгвізму з точки зору соціолінгвістики важливо також згадати фактори, які впливають на намір вивчення другої мови. Насамперед це мотивація, що розглядається як сукупність причин, що викликають у людини зацікавленість у вивченні нової мови. У зв'язку з цим

виділяють два види мотивації: інструментальну та інтегративну. Термін «інструментальна мотивація» означає намір вивчати другу мову з метою отримання певних переваг, таких як підвищення власного професійного статусу, вдосконалення в навчанні, доступ до науково-технічної інформації тощо; у цьому випадку мова розглядається як інструмент досягнення прагматичних цілей. Термін «інтегративна мотивація» стосується інтересу чи бажання наблизитися до культури іншої мови та навіть інтегрувати її у власну культуру мовця [44, с.375].

Таким чином, до концепції білінгвізму можна підходити з позицій різних дисциплін і описувати її з різних теоретичних перспектив. Психолінгвістика розуміється як наука, яка аналізує вплив білінгвізму на ментальні та когнітивні процеси людини, тоді як соціолінгвістика вивчає вплив двомовності на спільноту чи суспільство, культурні норми та фактори, що сприяють вивченню другої мови людиною. Білінгвізм як складне явище враховує не лише лінгвістичну перспективу, а й психологічні, соціальні та культурні аспекти поряд із нею.

## **Висновки до розділу 1**

Білінгвізм є досить складним і неоднозначним явищем, яке не має консенсусу, тому різні теоретичні дисципліни та міждисциплінарні напрями розробили безліч його описів і тлумачень протягом останнього століття. При дефініції терміну білінгвізм враховуються різні критерії, головним з яких є практика почергового використання двох мов, що передбачає можливість їхнього застосування у комунікативних цілях. Важливим для лінгвістики є також рівень володіння мовами, який має бути достатнім для невимушеної розмови та висловлення власних думок.

У даний час існує кілька теоретичних перспектив та напрямів вивчення білінгвізму, серед яких особливий інтерес викликають психолінгвістичний і соціолінгвістичний, а також власне лінгвістичний. Лінгвістика сама по собі має справу з білінгвізмом, пов'язаним із лінгвістичною інтерференцією та мовними

контактами, тоді як у психолінгвістиці основна увага приділяється процесам сприйняття та розуміння мов, характеру їх взаємодії в мисленні і мовленнєвій діяльності білінгва. Соціолінгвістика двомовності аналізує різні варіації між соціальною та мовною структурами, визначає фактори функціонування білінгва в суспільстві поряд із визначенням його намірів та сфери застосування кожної з двох мов. Таким чином, білінгвізм як феномен комплексного характеру враховує не лише лінгвістичну перспективу, а й психологічні, соціальні та культурні аспекти поряд із нею.

## РОЗДІЛ 2. ЛІНГВІСТИЧНА ІНТЕРФЕРЕНЦІЯ В УМОВАХ БІЛІНГВІЗМУ

### 2.1. Загальне поняття інтерференції, її причини та прояви на різних мовних рівнях

Білінгвізм не є унікальним феноменом, але з точки зору людських потреб та можливостей комунікації це суспільне і вкрай неоцінене явище веде до найрізноманітніших мовних контактів, які руйнують як територіальні, так і соціальні межі. Численні дослідження білінгвізму та, як його наслідок, мовних взаємодій (Аканде і Акінвал (2006), Кристал (1997), Кук (1992, 2002), Грожан (1989), Хауген (1953, 1956) та Вайнрайх (1974)) показують, що співіснування двох мов у свідомості індивіда є складним явищем, яке передбачає зовсім інше розуміння мови білінгвом порівняно з монолінгвом. Як зазначає у своїй праці Аменорві [17], використання людиною двох і більше мов включає такі фактори, як рівень володіння мовами, їхнє призначення (з якою метою людина використовує різні мови, які ролі вони відіграють у загальній моделі поведінки індивіда), почергове використання (ступінь чергування двох мов, спосіб переходу з однієї мови на іншу та за яких умов це відбувається), а також інтерференцію – наскільки білінгв розмежовує чи, навпаки, об'єднує різні мови, як особливості однієї мови впливають на процес використання іншої тощо.

Мовна інтерференція – об'єктивний процес, який відбувається під час мовного контакту, незалежно від того бажає цього людина чи ні. Зазвичай дане явище «працює» в одному напрямку й означає відхилення від норм системи другої мови під впливом особливостей рідної. Проте водночас мовна інтерференція явище динамічне, яке, як правило, послаблюється в процесі оволодіння іноземною мовою [32, с.273].

Марлін Двінастіні припускає, що білінгвізм є однією з основних причин мовних взаємодій і розглядає інтерференцію як негативний аспект двомовності.

Він стверджує, що коли розуміння людиною однієї мови позначається на її розумінні іншої мови, тоді ця людина зазнає мовного впливу. Він може бути негативним, по-іншому відомим як інтерференція, коли розуміння однієї мови ускладнює розуміння іншої мови. Як альтернатива, може бути позитивний вплив, коли знання однієї мови є корисним при вивченні другої мови. Чим більші відмінності між двома мовами, тим більш негативними будуть наслідки інтерференції. Це неминуче трапиться в ситуації, коли хтось повністю не опанував другу мову [27].

Термін «інтерференція» має ряд тлумачень у науці. У сфері лінгвістики широкий діапазон визначень інтерференції можна згрупувати за чотирма типами. Інтерференцію можна розглядати як: 1) небажане проникнення особливостей рідної мови в іноземну (Кулешов, Харченко); 2) взаємодію структур і елементів структур двох мовних систем як результат мовних контактів (Хауген, Жлуктенко); 3) явище, яке відбувається в системі й структурі двох мов і має системний характер (Вайнрайх); 4) випадки відхилень мовлення білінгвів у порівнянні з монолінгвами (Хауген, Вайнрайх) [32].

Загалом лінгвістична інтерференція охоплює всі рівні мовної системи: від фонологічного до синтаксичного. У фонології інтерференція виявляється в неправильній інтонації, побудові звуків, іншомовному акценті тощо. Граматична інтерференція має прояв у неправильному вживанні артиклів, порядку слів, отоженні слів, граматичних моделей мов [10, с.58]. Що ж до лексичної інтерференції, її причинами є семантичні й структурні розбіжності в лексичних системах двох мов, які проявляються в запозиченнях, неправильному перекладі слів, які є майже ідентичними з морфологічної точки зору, однак мають абсолютно різний переклад. Синтаксична інтерференція пов'язана з неправильною побудовою синтаксичних конструкцій, порушенням синтаксичних зв'язків між словами під впливом безпосереднього контакту синтаксичних систем двох мов [1, с.23].

Ступінь виявлення інтерференції на різних мовних рівнях зумовлений ступенем відмінностей між двома мовами, іншими словами – їх типологічною

близькістю або розрізненням. Під час вивчення другої мови надзвичайно важливо враховувати особливості її фонетичної системи, зокрема і просодичної підсистеми, адже міжмовна інтерференція особливо яскраво спостерігається саме під час вимови. Наприклад, люди, які вивчають іноземну мову, краще засвоюють артикуляцію того звуку мови, якого немає в їх рідній, ніж артикуляцію звуку, який є алофоном у їх рідній мові [34]. Однак, чим більше схожі фонологічні системи двох мов, тим легше освоїти другу мову [32]. У цьому відношенні потенційний обсяг перешкод для вивчення другої мови прямо пропорційний кількості відмінностей між двома мовами.

## **2.2. Особливості українсько-англійської просодичної інтерференції**

На думку У.Вайнрайха, особливо важливою є інтерференція, яка відбувається на мовленнєвому рівні та визначається як «випадки відхилення від норм однієї з мов, що виникають у мовленні білінгвів внаслідок їхнього володіння більш ніж однією мовою, тобто в результаті мовного контакту» [47, с.2]. Інтерференція у мовленні білінгва охоплює різні рівні системи мови, від фонологічного до синтаксичного, проте однією з найменш досліджуваних й водночас найбільш вагомих є просодична інтерференція, яка виявляється у відхиленнях від просодичних норм мови, що є наслідком впливу іншої мови.

З усіх видів інтерференції (фонетичної, морфологічної, граматичної, лексичної – семантичної чи словотвірної, синтаксичної) просодична є найбільш звичною та розповсюдженою; дуже часто особливості просодичної інтерференції виникають у мовленні білінгва, навіть якщо він добре володіє другою мовою.

Підкреслюючи роль інтонації в мовленні, Р. Кінгдон писав: «Інтонація – це душа мови, тоді як вимова її звуків – її тіло. Існує практична причина, чому доцільно приділяти більше уваги інтонації, ніж вимові. Звуки англійської мови, як вони вимовляються різними мовцями та на різних діалектах, варіюються в широких межах, так що іноземець, який вивчає англійську мову, має певну свободу в цьому плані, але на більшості діалектів наголос та інтонація досить

тісно відповідають одній і тій же схемі» [31, с.40]. Ми не можемо недооцінювати значення явища просодичної інтерференції, особливо в інтонованих мовах, таких як українська та англійська. Кожна з них має свої інтонаційні контури, що допомагають мовцю передати власні емоції, ставлення до співрозмовника чи зрозуміти прагматичне значення слів інших. Важливо враховувати, яким чином інтонаційні моделі української мови можуть впливати на англійську та сам процес спілкування.

Опис явища просодичної інтерференції є необхідним, для того, щоб оцінити, як особливості української мови впливають на процес мовлення англійською. Різні дослідження українсько-англійської просодичної інтерференції [5; 6], показують, відхилення від норм просодичної структури другої мови спостерігається у всіх підсистемах на слуховому та акустичному рівнях, зокрема ЧОТ, наголос, ритм і темп є основними зонами цих відхилень.

Фонетичним терміном на позначення висоти звуку є частота основного тону (ЧОТ), яка визначається частотою вібрацій голосових зв'язок. Коли кількість вібрацій у секунду (ЧОТ) зростає, висота звуку також зростає, хоча їхнє співвідношення не прямо пропорційне одне одному [41]. В англійській мові існують чіткі інтонаційні моделі. Наприклад, у межах однієї інтонаційної групи в англійській мові зустрічаємо ширший інтервал між високим та низьким тонами у порівнянні з українською, а зниження чи підняття інтонації в кінці речення є виразнішим і несе істотніше смислове навантаження. До того ж висхідна чи спадна інтонація в англійській мові може мати свою соціальну практику. В українській ж мові, речення, у яких висловлюється почуття жалю чи якесь прохання, зазвичай вимовляються з низхідним тоном, а не висхідним (як в англійській мові) [3]. Порівняйте: «I'm very sorry to have been late» (загальне підняття інтонації до кінця речення) та «Мені дуже шкода, що я запізнився» (зниження інтонації).

В українській мові висота голосу є більш монотонною, тобто немає різких контрастів (переходів) у ЧОТ між структурними елементами висловлювання, а через вплив високих частот рідної мови українці часто використовують

інтонаційні моделі, не характерні для англійської мови, наприклад висхідний тон замість низхідного [8, с.80].

Ритм та наголос є областями найбільшої кількості відхилень від просодичної норми нерідної мови білінгва [41, с.53]. Очевидно, що це твердження впливає із функціональних та структурних відмінностей між просодичними (тонічними) підсистемами фонетичних систем мов, що контактують. Особливостями міжмовної інтерференції в англійському мовленні українських білінгвів є:

1. Більша кількість акцентних ритмів та як наслідок спотворення ритмічних моделей англійської мови.
2. Зсув ядра речення вліво, на початку речення, як характерної риси для української мови.
3. Перебільшена виразність ядерного складу, що досягається за рахунок ширшого інтервалу кінцевого тону і більшої кількісної контрастності з ненаголошеними складами.
4. Перебільшена виразність наголошених складів у доядерній частині речення в результаті їх більшої тривалості, а також завдяки впливу контрастів ЧОТ.
5. Спотворення у відчутній ізохронії ритмічних одиниць під впливом складової часової тенденції українського ритму. Хоча українська мова відноситься до ізоакцентних як і англійська, тобто таких, що виявляють рівномірність інтервалів між наголошеними складами у мовленні, варіативність тривалості ритмічної одиниці в англійському мовленні українців більша, ніж у мовленні носіїв англійської мови. Розгорнуте англійське мовлення для ненаренованого «іноземного» вуха видається настільки ритмічним, що інколи нагадує віршоване [41, с.53].

Варто також зазначити, що в англійському мовленні службові (допоміжні) слова (прийменники, частки, займенники та ін.) не наголошуються, у той час як наголошеними залишаються слова з предметним чи поняттєвим значенням [3]. Порівняйте:

- Will you buy these flowers for mom? – наголос падає на слова will, buy, flowers, mom.
- Ти купиш мамі ці квіти? – наголошеними є усі слова [8, с.80].

Темп мовлення: згідно з дослідженнями просодичних особливостей систем української та англійської мов [8; 41], темп мовлення носіїв англійської мови є набагато вищим, ніж темп українських білінгвів при мовленні англійською. Це зумовлено безпосереднім впливом характерного для української мови низького темпу мовлення та рядом інших причин, серед яких:

1. Українські білінгви при мовленні англійською не скорочують ненаголошені склади так як це роблять носії мови.
2. Українські білінгви при мовленні англійською подовжують наголошені склади як і в рідній мові [8, с.80].

Усі вище перераховані відхилення від просодичних норм системи англійської мови проявляються у наявності українського акценту в білінгвів. Очевидно, що акцент – складне просодичне явище, на яке окрім артикуляційних особливостей контактуючих мов, впливає і їхня акцентуальна та ритмічна організація, закономірності наголошування, структура інтонаційних груп, правила розподілу речень на синтагми, специфічні інтонаційні ознаки тощо [32, с.275].

Варто також зазначити, що дослідження в галузі мовної інтерференції в цілому та просодичної, зокрема, є цінним матеріалом для теорії мовних контактів, оскільки виявляють закономірності у функціонуванні контактуючих мовних систем. Це також важливо з боку практичного застосування отриманих даних. Для вчителя іноземної мови, основною метою якого є підготовка своїх учнів до умов штучного білінгвізму, дуже важливо враховувати особливості просодичних систем контактуючих мов. При вивченні англійської мови варто пам'ятати як про особливості форм, так і про значення англійської просодії, тобто виробити правильні просодичні зразки та використовувати їх у відповідній ситуації.

## Висновки до розділу 2

Білінгвізм веде до численних мовних контактів, до яких належить і лінгвістична інтерференція. Мовна інтерференція – об'єктивний процес, який означає відхилення від мовних норм системи другої мови під впливом особливостей рідної. Водночас це явище динамічне, яке зазвичай послаблюється в процесі оволодіння іноземною мовою.

Лінгвістична інтерференція охоплює різні рівні системи мови, від фонологічного до синтаксичного. Ступінь виявлення інтерференції на різних мовних рівнях зумовлений ступенем відмінностей між двома мовами, іншими словами – їх типологічною близькістю або розрізненням. Таким чином потенційний обсяг перешкод для вивчення другої мови прямо пропорційний кількості відмінностей між двома мовами.

Особливо важливим та необхідним є аналіз явища просодичної інтерференції, яка спостерігається у зміні контурів ЧОТ, наголосу, ритму й темпу. Дослідження інтерференції на рівні просодики є цінним матеріалом для теорії мовних контактів, оскільки виявляють закономірності чи розбіжності у функціонуванні контактуючих мовних систем. Це також важливо з боку практичного застосування отриманих даних.

## РОЗДІЛ 3. СТВОРЕННЯ ТА АВТОМАТИЧНИЙ АНАЛІЗ КОРПУСУ МОВЛЕННЯ БІЛІНГВІВ

### 3.1. Створення корпусу мовлення білінгвів

Для дослідження просодичних особливостей мовлення білінгвів ми вирішили створити акустичний корпус. У студійних умовах у лабораторії експериментальної фонетики Інституту філології було зроблено записи читаного мовлення трьох дикторів – Троїцької Валерії, Гришина Іллі та Приймак Оксани – диктори 1, 2 та 3 відповідно. Решта записів були зроблені в домашніх умовах за допомогою диктофонів на мобільних пристроях. Для читаного мовлення кожного диктора (окрім диктора 6 – Іванишина Назара) були підготовлені тексти із сучасної художньої прози, а саме уривки із романів “Гаррі Поттер і філософський камінь” та “Гаррі Поттер і таємна кімната” англійської письменниці Джоан Роулінг, відібрані з мережі Інтернет. Окрім підготовленого тексту, для шести дикторів було запропоновано записати власне спонтанне мовлення англійською та українською мовами. Загальний обсяг зібраного матеріалу складає 2 год 10 хв 25 с (28 аудіозаписів); тривалість записів спонтанного мовлення становить 1 год 13 хв 24 с (12 аудіозаписів), читаного – 57 хв 1 с (16 аудіозаписів).

У дослідженні загалом взяли участь дев'ять дикторів: шість жінок та троє чоловіків. Це люди різного віку, професій, з різних регіонів України та світу. Проте усі диктори вважаються білінгвами, так як з народження вони є носіями української мови, проте у повсякденному житті, під час роботи чи навчання, у комунікативних цілях також використовують англійську мову, якою володіють вільно або на високому рівні (див. Додаток 1).

У лабораторії експериментальної фонетики записи мовлення дикторів здійснювалися за допомогою мікрофона «Shure BG5» та програмного забезпечення «Sound Forge Pro» (10.0); у домашніх умовах – на диктофони смартфонів різних марок та моделей. Середня тривалість кожного запису

становить від 4 до 9 хвилин (див. табл. ). Уривки з книг про Гаррі Поттера (англійською та відповідні їм українською мовами), середня кількість слів яких становить близько 600, диктори читали безперервно, з попередньою підготовкою. Для спонтанного мовлення мовці могли обрати або власну тему, або тему із заздалегідь підготовленого списку (див. Додаток 2). У більшості випадків диктори давали досить розгорнуті відповіді, ділились своїми думками, описували ситуації з життя, розповідали про себе, свої хобі та зацікавлення.

Диктор	Тривалість записів ЧМ укр. (хв)	Тривалість записів ЧМ англ. (хв)	Тривалість записів СМ укр. (хв)	Тривалість записів СМ англ. (хв)
Диктор 1	3:43	3:34	6:05	8:28
Диктор 2	4:24	4:25	7:58	9:09
Диктор 3	4:07	5:37	—	—
Диктор 4	3:14	3:43	—	—
Диктор 5	2:47	3:43	5:01	4:52
Диктор 6	—	—	4:35	5:19
Диктор 7	3:55	4:42	—	—
Диктор 8	4:04	4:19	6:24	4:55
Диктор 9	4:28	4:24	5:15	5:56

Таблиця 3.1

### ЗАГАЛЬНА ТРИВАЛІСТЬ ЗАПИСІВ

Для створення корпусу мовлення використовувалася комп'ютерна програма Praat [42]. Praat – це безкоштовний комп'ютерний пакет програм для фонетичного наукового аналізу мовлення. Praat був розроблений Полом Боерсма і Девідом Венінк з Амстердамського університету, сьогодні виходять нові версії, Praat продовжує розвиватися. Він може працювати на широкому діапазоні операційних систем, включаючи різні версії Unix, Linux, Mac і Microsoft Windows. Програма також підтримує синтез мови, включаючи артикуляційний синтез.

У програмі Praat ми аналізували записи з допомогою анотації у форматі TextGrid, де файл було проанатовано та просегментовано. Анотація файлів виконувалася на чотирьох рівнях:

- 1) синтагма;
- 2) примітки, де вказувалися певні індивідуальні особливості вимови диктора: дефекти мовлення, дихання тощо);
- 3) частота основного тону (ЧОТ; для кожної синтагми вираховувалася середня ЧОТ);
- 4) інтенсивність (як і у випадку з ЧОТ, для кожної синтагми визначалася середня інтенсивність).

### **3.2. Синтагматичне членування записів мовлення**

Як уже було сказано, кожен звуковий запис був поділений на синтагми. Термін «синтагма» використовуємо в значенні мінімального інтонаційно-сміслового та структурного компонента мовленнєвого потоку (як читаного, так і спонтанного мовлення) , який складається з одного або більше слів, пов'язаних структурою й за змістом і характеризується типовістю інтонаційного контуру. Синтагматичне членування мовного потоку є виявом членування думки на окремі, в певний спосіб пов'язані між собою смислові комплекси, тобто синтагми. Основою синтагматичного членування слід вважати завершеність поняття, що виражається синтагмою, і його здатність відображати реальну дійсність.

Членування звукових записів відбувалось з урахуванням особливостей членування потоку мовлення самим диктором:

1. Основний критерій – визначили межі синтагми за паузами. Варто зазначити, що до уваги бралися як логічні (змістові), так і психологічні паузи. Психологічна пауза, на відміну від логічної, що служить для змістового виділення слова чи синтагми або ж вираження характеру зв'язку між частинами висловлювання, є важливим засобом передачі емоцій, настрою і бажань мовця. Зазвичай психологічну паузу

використовують, щоб зосередити увагу співрозмовника на значенні слів, що йтимуть після неї, або ж дати можливість слухачам сприйняти, пережити вже сказані слова. Були взяті до уваги й хезитаційні паузи, які особливо яскраво використовуються дикторами під час спонтанного мовлення для розмірковування (обдумування).

2. Визначили індивідуальні особливості вимови диктора (дефекти мовлення, дихання, наявність подовження голосних та приголосних, довгі паузи між синтагмами, хезитації і тд.).
3. Застосовували семантичний критерій (враховували контекст). Тобто розглядали синтагму як одну, що виражала у певному контексті одне поняття.

У ході дослідження встановлено, що фразові реалізації українського та англійського варіантів протиставляються за: 1) характером членування на синтагми; 2) наявністю синтагм різної складової протяжності; 3) тривалості пауз між синтагмами.

### 3.3. Створення скриптів для аналізу просодики

Для автоматичного аналізу просодичних особливостей мовлення дикторів нами були розроблені скрипти у програмі Praat. Скрипт – текст (програмний код), який автоматизує послідовність операцій під час аналізу звукових файлів. Для кожного аудіозапису було запущено кілька скриптів у певному порядку:

1. *boundaries\_lv3-4.praat* – скрипт автоматично розставляє ті самі межі синтагм, що знаходяться на рівні 1 (phrases) у TextGrid-файлі, на рівнях 3 (pitch) і 4 (intense) (див. Додаток 3).
2. *meanpitch.praat* – скрипт автоматично визначає середній показник ЧОТ для кожного інтервалу рівня pitch, якщо відповідний йому інтервал на рівні phrase заповнений текстом (тобто не є паузою) (див. Додаток 4).
3. *meanintensity.praat* – скрипт автоматично визначає середній показник ІЗС для кожного інтервалу рівня intense, якщо відповідний йому інтервал на рівні phrase заповнений текстом (див. Додаток 5).

4. *mean\_all.praat* – скрипт автоматично вираховує середнє значення всіх середніх значень ЧОТ та інтенсивності (див. Додаток 6).
5. *count\_pauses.praat* – скрипт автоматично підраховує кількість пауз (окремо коротких <150ms, середніх 150-300ms і довгих >300ms) у розміченому звуковому файлі. Паузами вважаються пусті інтервали на рівні 1 (див. Додаток 7).
6. *get\_phrases.praat* – скрипт отримує з файлу *.TextGrid* усі синтагми і зберігає їх у новому текстовому файлі – кожен синтагму на новому рядку. (див. Додаток 8).

Перед запуском скриптів потрібно у вікні завантажених файлів програми Praat виділити звук і відповідний йому TextGrid-файл або лише TextGrid-файл (для скрипта *boundaries\_lv3-4.praat*).

### 3.3.1. Скрипт *boundaries\_lv3-4.praat*

Скрипт *boundaries\_lv3-4.praat* має такий алгоритм:

1. Перевірка наявності файлів.

Перевіряє, чи користувач виділив файл TextGrid у вікні Praat: `numberOfSelected ("TextGrid")` і повертає повідомлення про помилку, якщо ні.

2. Виділення об'єктів (команди `selected()`, `selectObject`).

3. Створення меж фраз на рівнях 3 і 4.

Алгоритм знаходить кількість інтервалів на рівні 1. Далі він заходить у цикл. У ньому для кожного інтервалу, крім першого, виконуються такі дії:

- отримати час початку інтервалу `phrase_start` (тобто час, на якому знаходиться межа) (команда `Get start time of interval`);
- встановити межу на рівні 3 у точці `phrase_start` (команда `Insert boundary`);
- встановити межу на рівні 4 у точці `phrase_start` (команда `Insert boundary`).

### 3.3.2. Скрипти *meanpitch.praat* та *meanintensity.praat*

Скрипти *meanpitch.praat* і *meanintensity.praat* мають однакову структуру (алгоритм), але обчислюють різні показники на різних рівнях файлу TextGrid. Їх загальний алгоритм:

#### 1. Перевірка наявності файлів.

Спочатку алгоритм перевіряє чи користувач виділив потрібні файли в вікні Praat. Якщо виявилось, що ні, алгоритм (за допомогою команди `exitScript`) завершує роботу скрипта й перед користувачем з'являється вікно з помилкою та повідомленням про необхідність виділення потрібних файлів.

#### 2. Виділення об'єктів.

У мові скриптів Praat у самих командах обробки інформації в файлах не зазначається конкретний об'єкт, до якого потрібно застосувати команду. Його необхідно "виділити" перед необхідними командами за допомогою функції `selected()` або команди `selectObject`. Саме це й відбувається в наступних рядках скрипта.

#### 3. Визначення діапазону ЧОТ/інтенсивності.

У Praat за замовчуванням завжди визначені числові межі акустичних показників мовлення, які використовуються при відображенні цих показників в інтерфейсі для роботи зі звуковими файлами. Але в скриптах Praat потрібно експліцитно зазначити межі показників, з якими він буде працювати.

Для ЧОТ межі визначаються за допомогою команди `To Pitch t, min, max`, де `t` – крок (між точками в часі для визначення ЧОТ), `min` – мінімальний показник ЧОТ, `max` – максимальний показник ЧОТ [46].

Для інтенсивності межі визначаються за допомогою команди `To Intensity min, t`, де `min` – мінімальний показник ЧОТ, `t` – крок (між точками в часі для визначення інтенсивності) [45].

#### 4. Визначення показників ЧОТ/інтенсивності у файлі TextGrid.

Алгоритм знаходить кількість інтервалів на рівні 1 (рівносильно рівням 3 й 4, адже кількість інтервалів на них однакова). Після цього за допомогою циклу

одночасно ітерує через кожен інтервал двох потрібних рівнів (1 і 3 або 1 і 4). Для кожного інтервалу окремо виконує такі дії, якщо інтервал рівня 1 не пустий (у ньому є текст):

- отримати часову точку початку інтервалу `t1` (команда `Get starting point`, значення зберігається у змінну);
- отримати часову точку кінця інтервалу `t2` (команда `Get end point`, значення зберігається у змінну);
- визначити середнє значення інтенсивності між точками `t1` і `t2` в дБ (команда `Get mean`);
- записати отримане значення в потрібний інтервал (команда `Set interval text`), переконвертувавши його в текст та округливши значення до цілих.

### 3.3.3. Скрипт `mean_all.praat`

Алгоритм скрипта `mean_all.praat`:

1. Перевіряє, чи користувач виділив файл `TextGrid` у вікні Praat:  
`numberOfSelected ("TextGrid")`
2. Вибирає файл `TextGrid`: `selected (), selectObject`
3. Створює змінні для суми значень ЧОТ, суми значень інтенсивності й кількості синтагм:  
`pitch_sum = 0`  
`intensity_sum = 0`  
`phrases = 0`
4. Витягує загальну кількість інтервалів на рівні 1:  
`n = Get number of intervals: 1`
5. За допомогою циклу ітерує через інтервали рівня 1: `for i to n`
6. Якщо інтервал не пустий (тобто фраза):
  - алгоритм отримує текст для цього інтервалу з рівня 3 (ЧОТ) і 4 (інтенсивність):  
`pitch$ = Get label of interval: 3, i`  
`intensity$ = Get label of interval: 4, i`
  - конвертує їх зі змінної типу "рядок" у змінну типу "число";

- ```
pitch = number (pitch$)
intensity = number (intensity$)
```
- додає отримані числа до відповідної змінної з сумою:

```
pitch_sum += pitch
intensity_sum += intensity
```
  - збільшує кількість синтагм на 1: `phrases += 1`
7. Після закінчення циклу алгоритм обчислює середні значення ЧОТ та інтенсивності: `mean_pitch = pitch_sum / phrases ; mean_intensity = intensity_sum / phrases`
8. Виводить їх на екран у файл Praat Info у форматі:
- ```
Середня ЧОТ: n
Середня інтенсивність: m
```

### 3.3.4. Скрипт `count_pauses.praat`

Алгоритм скрипта *count\_pauses.praat*:

1. Перевірка наявності одного виділеного файлу *.TextGrid* у вікні Праату:

```
numberOfSelected ("TextGrid")
```
2. Вибір файлу *.TextGrid*: `selected (), selectObject`
3. Створення окремої змінної для обчислення кількості кожного типу пауз. Кожній змінній на початку присвоюється значення 0:

```
short = 0 ; medium = 0 ; long = 0
```
4. Отримання з файлу *.TextGrid* кількості інтервалів на рівні 1. За допомогою циклу програма ітерує через інтервали рівня 1:

```
n = Get number of intervals: 1
  for i to n
```
5. Цикл: для кожного інтервалу рівня 1 виконуються такі дії:
  - якщо інтервал пустий (тобто позначає паузу), алгоритм отримує часові точки початку й кінця інтервалу й на їх основі обчислює тривалість паузи в секундах:

```
if label$ = ""
  ts = Get start point: 1, i
  te = Get end point: 1, i
```

```
pause_length = te - ts
```

- якщо тривалість паузи  $< 150ms$ , одиниця додається в змінну *short*.  
Якщо тривалість паузи  $150-300ms$ , одиниця додається в змінну *med*.  
Якщо тривалість паузи  $> 300ms$ , одиниця додається до змінної *long*.

```
if pause_length < 0.150
    short += 1
elseif pause_length > 0.3
    long += 1
else
    medium += 1
```

6. Після закінчення роботи циклу, алгоритм обчислює загальну кількість пауз і записує в відповідний файл *.csv* дані відповідно до структури таблиці. Додаткові обчислення (відсоткова частка) і конвертування чисел у рядки відбуваються в цій самій команді.

### 3.3.5. Скрипт *get\_phrases.praat*

Алгоритм роботи скрипта:

1. Перевіряє, чи користувач виділив файл *.TextGrid* у вікні Праату:  

```
numberOfSelected ("TextGrid")
```
2. Вибирає файл *.TextGrid*: 

```
selected (), selectObject
```
3. Отримує загальну кількість інтервалів на рівні 1:  

```
n = Get number of intervals: 1
```
4. Зберігає у змінну назву файлу *.TextGrid*, аби створити текстовий файл із такою самою назвою:  

```
objName$ = selected$: "TextGrid"
```
5. За допомогою циклу ітерує через інтервали рівня 1: 

```
for i to n
```
6. Якщо інтервал не пустий (позначає синтагму, а не паузу), вставляє в новий рядок текстового файлу текст синтагми:  

```
appendFileLine: "./phrases_txt/" + objName$ + ".txt", label$
```

### 3.4. Розробка програмних кодів для аналізу синтагм

Окрім скриптів, нами було розроблено кілька програмних кодів для автоматичного аналізу довжини синтагм у читаному та спонтанному мовленні дикторів:

1. *edit\_phrases.py* – програма з'єднує в синтагмах частини фонетичних слів, тобто прибирає пробіл між ними, для зручності подальшого автоматичного обчислення кількості слів у синтагмах (див. Додаток 9).
2. *count\_phrases.py* – програма обчислює кількість кожного типу синтагм (однослівні, двослівні і т.д.) у текстовому файлі й зберігає інформацію у файлах формату .csv (див. Додаток 10).

#### 3.4.1. Програма *edit\_phrases.py*

Алгоритм програми має такі етапи роботи:

1. Визначення шляху до папки, у якій зберігаються текстові файли з синтагмами, і папки, у якій програма зберігатиме нові текстові файли з відредагованими синтагмами:

```
path = r'.\phrases_txt'  
new_path = r'.\phrases_edited'
```

2. Створення регулярних виразів для редагування синтагм. Регулярних виразів два: один для англійської мови, інший – для української. Обидва регулярні вирази шукають у синтагмі патерн «слово довжиною *n* символів + пробіл + наступне слово» і прибирають у ньому пробіл, створюючи фонетичне слово. Для англійської мови *n*, для української мови *n*. Значення *n* для обох мов обиралося на власний розсуд у результаті експериментів, у ході яких визначали, із якими значеннями найбільш точно правильно визначалися фонетичні слова:

```
regex = {  
    True: [r"(\b[''\w-]{1,4}\b) (\b[''\w-]+)", r'\1\2'], # eng  
    False: [r"(\b[''\w-]{1,3}\b) (\b[''\w-]+)", r'\1\2'] # ukr  
}
```

3. Оскільки внаслідок роботи скрипта `get_phrases.praat` текстові файли з синтагмами були збережені з кодуванням Юнікод Big Endian, а в програмах мовою Python ми звикли працювати з кодуванням файлів UTF-8, необхідно було зберегти файли в цьому кодуванні. Зробили це автоматично:

```
with open(path + '\\\\' + filename, 'r', encoding='utf-16-be') as file:
    phrases = file.read()
with open(path + '\\\\' + filename, 'w', encoding='utf-8') as file:
    file.write(phrases)
```

4. Автоматичне визначення фонетичних слів із застосуванням регулярних виразів. Вміст кожного текстового файлу було поділено на синтагми, кожна синтагма була відредагована за допомогою потрібного регулярного виразу, відредаговані синтагми були збережені у новому файлі:

```
for n in range(len(phrases)):
    phrases[n] = phrases[n].strip('\n')
    if phrases[n] != '':
        phrases[n] = re.sub(regex['eng' in filename][0],
                            regex['eng' in filename][1],
                            phrases[n])
```

### 3.4.2. Програма `count_phrases.py`

Алгоритм програми має такі етапи роботи:

1. Визначення шляхів до файлів, з якими буде працювати алгоритм:

```
path = r'.\phrases_edited'
read_csv = open(r'.\phrases_read.csv', 'a', encoding='utf-8')
spontan_csv = open(r'.\phrases_spontan.csv', 'a', encoding='utf-8')
```

2. Створення списку `phrase_count`, за допомогою якого буде здійснюватися обчислення кожного типу синтагм окремо:

```
phrase_count = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
```

3. Зчитування вмісту текстових файлів із синтагмами, поділ вмісту на синтагми, поділ синтагм на фонетичні слова, додавання одиниці до

елементу списку `phrases_count`, індекс якого відповідає кількості фонетичних слів у синтагмі:

```
for n in range(len(phrases)):  
    try:  
        phrase_count[len(phrases[n].strip().split(' '))] += 1  
    except IndexError:  
        phrase_count[-1] += 1
```

4. Збереження потрібної інформації в заздалегідь створений файл `.csv`, відповідно до структури таблиці в ньому.

### Висновки до розділу 3

Створений у результаті дослідження корпус читаного та спонтанного мовлення білінгвів існує у формі сукупності звукових файлів і відповідних їм файлів анотації. Загальний обсяг зібраного матеріалу складає 2 год 10 хв 25 с (28 аудіозаписів); тривалість записів спонтанного мовлення становить 1 год 13 хв 24 с (12 аудіозаписів), читаного – 57 хв 1 с (16 аудіозаписів). Файли анотації створювалися за допомогою програми *Praat*. Було створено 4 рівні анотації:

- 1) синтагма;
- 2) примітки (індивідуальні особливості вимови диктора: дефекти мовлення, дихання тощо);
- 3) середня ЧОТ для синтагми;
- 4) середня інтенсивність для синтагми.

У процесі анотування корпусу, а також для подання фонетичних показників мовлення у зручній для аналізу формі, використовували можливості мови скриптів *Praat* і мови програмування *Python*.

Мовою *Praat* були створені такі скрипти: *boundaries\_lv3-4.praat* – розставляє ті самі межі синтагм, що знаходяться на рівні 1 (`phrases`) у *TextGrid*-файлі, на рівнях 3 (*pitch*) і 4 (*intense*); *meanpitch.praat* – визначає середній показник ЧОТ на рівні *pitch* для кожної синтагми; *meanintensity.praat* – визначає середній показник ІЗС на рівні *intense* для кожної синтагми; *mean\_all.praat* – обчислює середнє значення всіх значень ЧОТ та інтенсивності

на рівнях *pitch* та *intense*; *count\_pauses.praat* – підраховує кількість коротких, середніх і довгих пауз у розміченому файлі; *get\_phrases.praat* – отримує з файлу *.TextGrid* усі синтагми й зберігає їх у новому текстовому файлі.

Мовою *Python* було створене таке програмне забезпечення: *edit\_phrases.py* – з'єднує в синтагмах частини фонетичних слів; *count\_phrases.py* – обчислює кількість кожного типу синтагм (однослівні, двослівні і т.д.) у текстовому файлі.

Створені програмні засоби дозволили нам автоматизувати аналіз *TextGrid*-файлів й у такий спосіб пришвидшили й полегшили нашу роботу.

## РОЗДІЛ 4. ПРОСОДИЧНІ ХАРАКТЕРИСТИКИ МОВЛЕННЯ БІЛІНГВІВ

У рамках нашого дослідження розглядається контакт двох мов – української у порівнянні з англійською. Говорячи про спільні та відмінні типологічні ознаки контактуючих мов, слід зазначити, що обидві мови, якими володіють учасники експерименту, відносяться до «нетональних» мов, але до різних мовних груп: англійська – до групи германських мов, українська – до групи слов'янських, але до спільної індоєвропейської сім'ї. Таким чином, ми можемо говорити про генетично обумовлені типологічні спільні риси чи відмінності в системах двох мов.

### 4.1. Особливості паузації

Паузам належить важлива роль у просодичному оформленні синтагм. У читаному мовленні паузи здебільшого виконують розділову функцію, тобто відокремлюють синтаксичні одиниці в тексті, що зумовлено змістом та пунктуаційним оформленням речень. У спонтанному ж мовленні паузи здебільшого спричинені не синтаксичними, а семантичними факторами. Характерною особливістю спонтанного мовлення є наявність значної кількості психологічних пауз, що зумовлено їх емоційним та функціональним навантаженням: диктор використовує психологічні паузи як засіб вираження емоційної оцінки, зосередження уваги співрозмовника, увиразнення змісту сказаного, передачі настрою і бажань. Також для спонтанного (та рідше читаного) мовлення є характерними хезитаційні паузи як показник вагання, розмірковування, підбиранням потрібного слова тощо.

Для того, щоб порівняти тривалість пауз в українському й англійському мовленні кожного диктора, ми автоматично створили дві таблиці з кількостями кожного типу пауз в анотованих файлах: одну таблицю для читаного мовлення, другу – для спонтанного. Таблиці були створені у форматі `.csv`, адже він є найзручнішим табличним форматом для автоматичної обробки даних за допомогою скриптів Praat.

Отож, першим кроком було створення вручну файлів формату *.csv*, у яких у першому рядку записали через коми назви стовпчиків таблиць: *filename* (назва файлу *.TextGrid*), *short* (кількість коротких пауз), *med* (кількість середніх пауз), *long* (кількість довгих пауз), *total* (загальна кількість пауз у файлі), *short%* (частка коротких пауз у відсотках від загальної кількості), *med%* (частка середніх пауз), *long%* (частка довгих пауз).

Окрім абсолютної частоти типів пауз, таблиці також містять їхню відсоткову частку від загальної кількості пауз у файлі. Це було зроблено з метою забезпечення наочності показників, аби мінімізувати людський фактор при аналізі й зробити правильні, достовірні висновки на основі даних таблиць.

Після визначення структури таблиць у вікні програми Praat ми запустили скрипт *count\_pauses.praat*, який заповнює їх потрібною інформацією, й отримали такі результати:

filename	short	med	long	total	short%	med%	long%
aneng	4	8	76	88	5%	9%	86%
anukr	2	18	70	90	2%	20%	78%
greng	0	17	105	122	0%	14%	86%
grukr	4	21	99	124	3%	17%	80%
kpeng	3	20	86	109	3%	18%	79%
kpukr	6	27	77	110	5%	25%	70%
kzeng	5	12	63	80	6%	15%	79%
kzukr	6	3	37	46	13%	7%	80%
mneng	1	4	50	55	2%	7%	91%
mnukr	3	9	43	55	5%	16%	78%
mteng	2	15	64	81	2%	19%	79%
mtukr	8	25	54	87	9%	29%	62%
preng	18	19	137	174	10%	11%	79%
prukr	14	21	79	114	12%	18%	69%
treng	0	16	61	77	0%	21%	79%
trukr	1	17	54	72	1%	24%	75%

Таблиця 4.1.

#### ТАБЛИЦЯ КІЛЬКОСТІ ПАУЗ ДЛЯ ЧИТАНОГО МОВЛЕННЯ

filename	short	med	long	total	short%	med%	long%
grsp_eng	0	12	150	162	0%	7%	93%
grsp_ukr	6	26	149	181	3%	14%	82%
ivsp_eng	0	18	129	147	0%	12%	88%
ivsp_ukr	2	12	89	103	2%	12%	86%

kpsp eng	0	6	82	88	0%	7%	93%
kpsp ukr	2	21	111	134	1%	16%	83%
kzsp eng	0	4	59	63	0%	6%	94%
kzsp ukr	5	12	92	109	5%	11%	84%
mtsp eng	3	2	70	75	4%	3%	93%
mtsp ukr	11	26	82	119	9%	22%	69%
trsp eng	2	29	142	173	1%	17%	82%
trsp ukr	5	18	119	142	4%	13%	84%

Таблиця 4.2.

#### ТАБЛИЦЯ КІЛЬКОСТІ ПАУЗ ДЛЯ СПОНТАННОГО МОВЛЕННЯ

Рядки таблиць слід розглядати попарно: кожен рядок відповідає одному диктору. При цьому перший із двох відповідає англійському мовленню, другий – українському.

На основі таблиць можемо зробити такі спостереження:

1. *Читане мовлення* (див. табл. 4.1):

- у кожного диктора, окрім диктора 5 (kz), в англійському мовленні порівняно з українським кількість довгих пауз більша на 4-17% (окремо маємо такі показники різниці: 8%, 6%, 9%, 13%, 17%, 10%, 4%). У диктора 5 ж кількість довгих пауз менша на 1%, що ми не вважаємо вагомим показником, адже збільшення тривалості пауз в англійському мовленні в неї все-таки спостерігається: коротких пауз на 7% менше, середніх пауз – на 8% більше;
- у кожного диктора (окрім диктора 5, причини ми розглянули вище) кількість середніх пауз в англійському мовленні менша на 3-11% (окремі показники різниці: 11%, 3%, 7%, 9%, 10%, 7%, 3%), що спричинено збільшенням кількості довгих пауз;
- у кожного диктора, окрім одного – диктора 7 (an), в українському мовленні більша кількість коротких пауз на 1-7% (окремі показники різниці: 3 2%, 7%, 3%, 7%, 2%, 1%).

*Спонтанне мовлення* (див. табл. 4.2). Спостереження збігаються з тими, що ми отримали для читаного мовлення, хоча є певні відмінності:

- у кожного диктора, окрім диктора 1 (tr), в англійському мовленні порівняно з українським кількість довгих пауз більша на 2-24%. Порівнюючи окремі показники різниці для спонтанного мовлення (11%, 2%, 10%, 10%, 24%) з показниками читаного, бачимо, що для спонтанного мовлення у більшій частині дикторів кількість довгих пауз збільшується на більший показник (для 4 з 5 дикторів – різниця 10-24% порівняно з читаним мовленням, де для 3 із 7 дикторів – різниця 10-17%);
- у кожного диктора (окрім диктора 1 – tr та диктора 6 – iv) кількість середніх пауз в англійському мовленні менша на 5-19%, що спричинено збільшенням кількості довгих пауз (окремі показники різниці: 7%, 9%, 5%, 19%);
- у кожного диктора кількість коротких пауз в українському мовленні більша на 1-5%, ніж в англійському (окремі показники різниці: 3%, 2%, 1%, 5%, 5%, 3%).

Отримані спостереження можемо узагальнити таким чином:

1. Перехід дикторів на англійську мову під час читаного й спонтанного мовлення завжди супроводжується збільшенням тривалості пауз.
2. У більшості дикторів кількість довгих пауз збільшується за рахунок значного зменшення кількості середніх пауз і незначного зменшення кількості коротких пауз.
3. У деяких дикторів натомість збільшується кількість середніх пауз за рахунок зменшення кількості коротких пауз.
4. У спонтанному мовленні тривалість пауз збільшується більше, ніж у читаному.

Імовірно, такі відмінності між англійським та українським мовленням дикторів пов'язані з тим, що для кожного диктора англійська мова – не рідна (натомість друга або третя мова), тому англійське мовлення не супроводжується таким самим ступенем невимушеності. Натомість диктори більше уваги приділяють вимові, осмисленню семантики, зокрема в спонтанному мовленні – підбору правильних слів і синтаксичних конструкцій. Відмінності між читаним

і спонтанним мовленням імовірно пов'язані з тим, що диктори мають менший досвід у синтезі англійського мовлення, ніж у його аналізі.

## 4.2. Довжина синтагм

Синтагматичне членування є важливою просодичною характеристикою усного мовлення, адже як і паузи, які йому сприяють, воно пов'язане з мовленнєвим диханням, із структурою висловлювання, із акцентами в дискурсі.

У нашому дослідженні ми вирішили проаналізувати тривалість синтагм в українському й англійському мовленні, виявити чи є певні закономірності, диференційні й інтегральні ознаки в синтагматичному членуванні мовлення дикторами-білінгвами. Довжина синтагм вимірюється у фонетичних словах, тому алгоритм автоматичного укладення таблиць довжини синтагм був складнішим, ніж для тривалості пауз.

За допомогою скрипта *get\_phrases.praat* (див. Додаток 8) в окремий текстовий файл для кожного запису ми зберегли всі синтагми, які містилися у файлах *.TextGrid*. В отриманих текстових файлах кожна синтагма знаходиться з новогр рядка для зручності автоматичної обробки.

Оскільки у файлах *.TextGrid* міститься орфографічний запис тексту синтагм, просте обчислення кількості слів у них повернуло б кількість машинних слів (сукупностей графем від пробілу до пробілу), а не фонетичних слів. Тому ми створили алгоритм *edit\_phrases.py* (див. Додаток 9) мовою програмування Python (оскільки можливості скриптів у Прааті не дають можливості виконувати потрібні нам складні операції), який з'єднує в синтагмах машинні слова у фонетичні слова, припускаючи, що патерн «слово довжиною 1-3 (1-4) символи + пробіл + наступне слово» – фонетичне слово, а тому між частинами патерну слід прибрати пробіл. Відредаговані синтагми були збережені в нових текстових файлах тієї самої структури, що й попередні.

Обчислення фонетичних слів кожної синтагми відбувається за допомогою наступної програми *count\_phrases.py* (див. Додаток 10). Перед запуском цієї програми власноруч створили нові таблиці (файли *.csv*): одну для читаного

мовлення, іншу – для спонтанного. Таблиці мають таку структуру: *filename* (назва файлу *.TextGrid*), *one* (кількість однослівних синтагм), *two* (к-сть двослівних синтагм), *three* (к-сть трислівних синтагм), *four* (к-сть чотирислівних синтагм), *five* (к-сть п'ятислівних синтагм), *six* (к-сть шестислівних синтагм), *seven* (к-сть семислівних синтагм), *eight* (к-сть восьмислівних синтагм), *nine* (к-сть дев'ятислівних синтагм), *ten<=* (к-сть синтагм довжиною 10 і більше слів), *total* (загальна к-сть синтагм у файлі), *one%* (частка однослівних синтагм у відсотках від загальної к-сті), *two%* (частка двослівних синтагм), *three%* (частка трислівних синтагм), *four%* (частка сотирислівних синтагм), *five%* (частка п'ятислівних синтагм), *six%* (частка шестисліних синтагм), *seven%* (частка семислівних синтагм), *eight%* (частка восьмислівних синтагм), *nine%* (частка дев'ятислівних синтагм), *ten<=%* (частка синтагм довжиною 10 і більше слів).

Програма *count\_phrases.py* проводить необхідні обчислення й заповнює таблицю відповідною інформацією. Після цього для зручності аналізу показників додали в таблицю стовпчики *one-three%* (частка синтагм довжиною 1-3 слова), *four-six%* (частка синтагм довжиною 4-6 слів), *seven-nine%* (частка синтагм довжиною 7-9 слів). Необхідні числові значення в цих стовпчиках обчислюються автоматично за допомогою формул Google Таблиць.

У результаті автоматичної обробки файлів отримали такі дані:

<b>filename</b>	<b>total</b>	<b>one-three%</b>	<b>four-six%</b>	<b>seven-nine%</b>	<b>ten&lt;=%</b>
aneng	92	52	26	13	9
anukr	94	36	44	14	5
greng	140	67	20	12	1
grukr	145	61	33	4	1
kpeng	110	57	34	8	1
kpukr	117	59	35	8	1
kzeng	82	43	42	13	1
kzukur	51	20	34	26	22
mneng	55	23	25	25	27
mnukr	54	19	37	26	19
mteng	87	29	40	22	8
mtukur	101	29	47	22	3
preng	174	82	18	0	0
prukur	123	64	32	4	0

treng	79	26	46	17	11
trukr	82	34	31	22	15

Таблиця 4.3

ФРАГМЕНТ ТАБЛИЦІ З КІЛЬКІСТЮ СИНТАГМ ДЛЯ ЧИТАНОГО  
МОВЛЕННЯ (для повної таблиці див. Додаток 11)

filename	total	one-three%	four-six%	seven-nine%	ten<=%
grsp_eng	162	51	33	14	2
grsp_ukr	179	61	26	8	4
ivsp_eng	145	90	9	1	0
ivsp_ukr	102	68	24	7	2
kpsp_eng	87	41	35	17	6
kpsp_ukr	133	52	38	7	5
kzsp_eng	62	49	40	10	2
kzsp_ukr	108	72	22	6	1
mtsp_eng	74	32	29	27	11
mtsp_ukr	118	66	27	7	1
trsp_eng	172	39	34	16	11
trsp_ukr	141	50	28	12	9

Таблиця 4.4

ФРАГМЕНТ ТАБЛИЦІ З КІЛЬКІСТЮ СИНТАГМ ДЛЯ СПОНТАННОГО  
МОВЛЕННЯ (для повної таблиці див. Додаток 11)

Рядки таблиць слід розглядати попарно: кожен два рядки відповідають одному диктору. При цьому перший із двох відповідає англійському мовленню, другий – українському. Синтагми довжиною 1-3 слова вважатимемо короткими синтагмами, довжиною 4-6 слів – середніми, довжиною 7-9 слів – довгими, довжиною 10 і більше слів – наддовгими. Усього було проаналізовано 3069 синтагм у реалізації українських дикторів-білінгвів.

На основі таблиць можемо зробити такі спостереження:

1. *Читане мовлення* (див. табл. 4.3.). Загалом відмінності між мовами в відсоткових співвідношеннях кожної групи фраз хаотичні, закономірності виділити важко. Єдина закономірність, яку можемо спостерігати, – збільшення в англійському мовленні кількості коротких синтагм (у 5 з 8 дикторів – на 4-23%) і зменшення кількості середніх (у 6 з 8 дикторів – на

1-18%). Але ця закономірність може бути спричинена відмінностями в особливостях синтаксичних структур мов: в англійській мові дуже часто трапляються однослівні відокремлені обставини, натомість в українській мові це рідке явище.

2. *Спонтанне мовлення* (див. табл. 4.4.). Ситуація зі спонтанним мовленням така сама: закономірні відмінності між англійським та українським мовленням виділити важко. Тим не менш, в англійському мовленні спостерігаємо збільшення кількості довгих синтагм (у 5 з 6 дикторів – на 4-20%) і зменшення кількості коротких (у 5 з 6 дикторів – на 10-34%). Це може бути спричинено тим, що для кожного диктора англійська мова – нерідна. Через можливу невпевненість у своїх здібностях розмовляти англійською, вони намагаються більш детально продумувати кожен фразу перед тим, як її виголошувати.

Отже, аналіз результатів нашого дослідження показав, що довжина синтагм як у читаному, так і в спонтанному мовленні не залежить від мови, якою говорить диктор. У читаному мовленні вона залежить від структури самого тексту, а також від особливостей структури мови (незалежно від індивідуальних мовленнєвих характеристик диктора). У спонтанному мовленні мова може здійснювати вплив на довжину синтагм лише в деяких аспектах, але виявлена закономірність сумнівна. Загалом довжина синтагм залежить від індивідуальних мовленнєвих характеристик диктора, які не змінюються залежно від мови.

### **4.3. Частота основного тону (ЧОТ) та інтенсивність (ІЗС)**

На думку дослідників-лінгвістів, найбільш інформативними просодичними компонентами є висота та інтенсивність голосу, які вимірюються показниками середнього діапазону частоти основного тону (ЧОТ) та інтенсивності звукового сигналу (ІЗС). Результати здійсненого нами аудитивного аналізу сили звучання голосу та частотних характеристик мовлення дикторів англійською та українською мовою представлені в табл. 4.5., 4.6.

<b>filename</b>	<b>mean_pitch</b>	<b>mean_intensity</b>
aneng	99	61
anukr	96	60
greng	138	60
grukr	134	61
kpeng	229	66
kpukr	244	71
kzeng	226	69
kzukur	217	70
mneng	194	60
mnukr	203	63
mteng	217	48
mtukr	233	52
preng	228	61
prukr	223	58
treng	256	62
trukr	244	60

Таблиця 4.5

#### СЕРЕДНЯ ЧОТ ТА ІНТЕНСИВНІСТЬ ДЛЯ ЧИТАНОГО МОВЛЕННЯ

<b>filename</b>	<b>mean_pitch</b>	<b>mean_intensity</b>
grsp eng	123	60
grsp ukr	119	63
ivsp eng	100	65
ivsp ukr	104	63
kpsp eng	220	62
kpsp ukr	216	69
kzsp eng	216	60
kzsp ukr	217	64
mtsp eng	218	45
mtsp ukr	213	48
trsp eng	247	58
trsp ukr	240	62

Таблиця 4.6

#### СЕРЕДНЯ ЧОТ ТА ІНТЕНСИВНІСТЬ ДЛЯ СПОНТАННОГО МОВЛЕННЯ

Рядки таблиць слід розглядати попарно: кожен два рядки відповідають одному диктору. При цьому перший із двох відповідає англійському мовленню, другий – українському.

На основі таблиць можемо зробити такі спостереження:

*Читане мовлення* (див. табл. 4.5):

- у показниках середнього діапазону ЧОТ видно незначну різницю (ЧОТ зростає чи навпаки спадає) під час мовлення першою та другою мовами у всіх дев'яти дикторів, у межах 3-16 Гц. Цікавим є той факт, що у Диктора 1 (tr) та Диктора 9 (mt) спостерігається найвищий показник у відмінностях середнього діапазону ЧОТ між двома мовами – 12-16 Гц, хоча їхні інтонаційні контури є одними з найбільш плавних та рівномірних;
- динамічні риси висловлень в англійському та українському мовленні дикторів передаються на акустичному рівні інтенсивністю звукового сигналу, що аудитивно сприймається як гучність. Результати акустичного аналізу динамічних показників українського та англійського читаного мовлення засвідчують, що у трьох із восьми дикторів інтенсивність голосу зростає у межах 1-3 Дб при мовленні другою мовою, у решти же мовців інтенсивність звуково сигналу вища на 1-5 Дб при мовленні рідною мовою.

*Спонтанне мовлення* (див. табл. 4.6):

- закономірні відмінності в ЧОТ між англійським та українським спонтанним мовленням виділити важко. Тим не менш, в англійському мовленні спостерігаємо збільшення середньої ЧОТ у чотирьох із шести дикторів на 4-7 Гц. Цікавим є той факт, що кожен із цих дикторів володіє англійською мовою на рівні вищому, ніж В2. Через можливу впевненість у своїх здібностях розмовляти англійською, вони намагаються використовувати ширший діапазон і помірну гучність, дещо змінюючи інтонаційне оформлення синтагм;
- результати акустичного й аудитовного аналізу показників середньої інтенсивності українського й англійського спонтанного мовлення засвідчують, що в п'яти з шести дикторів (окрім, диктора 6 – iv) ІЗС зростає в межах 3-7 Дб при мовленні першою мовою. Ймовірно, така закономірність пов'язана з тим, що українська мова є рідною для всіх мовців і вони вільно володіють нею, тому підвищена гучність та широка

амплітуда коливань контуру інтенсивності надає їхньому мовленню більшого емоційного забарвлення, сприяє кращому розумінню висловленого й передачі прагматичного смислу.

Таким чином, аналіз результатів нашого дослідження показав, що ЧОТ як у читаному, так і в спонтанному мовленні майже не залежить від мови, якою говорить диктор. Це може бути пов'язано з тим, що діапазон частоти основного тону для кожного мовця залишається тим самим, а зміна інтонаційного оформлення висловлень залежить від особливостей структури самого тексту, а також від особливостей структури мови (незалежно від індивідуальних мовленнєвих характеристик диктора). У близько 70% (10 із 14) випадків, ІЗС дикторів зростає при мовленні (як читаному, так і спонтанному) першою мовою. Припускаємо, це зумовлено тим, що диктор почуває себе вільніше й більш обізнаним у знанні рідної мови, тому підвищена інтенсивність голосу сприяє емоційному насиченню й вираженому почуттю впевненості у власних словах.

#### 4.4. Темп мовлення

Темп мовлення – часова характеристика, яка поряд із висотою, силою й тембром голосу утворюють усне мовлення. Темп мовлення є виразником інтонування як цілої синтагми, так і окремого слова. Для прикладу, наголошені склади в слові вимовляються дещо повільніше, порівняно із ненаголошеними.

Лінгвістичні дослідження свідчать, що українська мова характеризується набагато нижчим темпом мовлення, порівняно з англійською. Це пов'язано як із особливостями самої мови (артикуляційні характеристики, звукова будова слів, фонетичні процеси – редуція, асиміляція, подовження тощо), так і з індивідуальними особливостями диктора (просодичні характеристики голосу, паузація й тд.).

Результати здійсненого нами аудитивного та акустичного аналізу темпу мовлення дикторів англійською та українською мовою представлені в табл. 4.7.

<b>filename</b>	<b>temp_read</b> (скл/с)	<b>filename</b>	<b>temp_spont</b> (скл/с)
-----------------	-----------------------------	-----------------	------------------------------

aneng	4		
anukr	5.5		
greng	4.3	grsp eng	4.3
grukr	5.2	grsp ukr	6.5
kpeng	4.4	kpsp eng	5.1
kpukr	6.0	kpsp ukr	6.4
kzeng	4.2	kzsp eng	3.5
kzukr	5.6	kzsp ukr	5.3
mneng	5.1		
mnukr	6.5		
mteng	5.1.	mtsp eng	4.3
mtukr	6.0.	mtsp ukr	5.1
preng	3.5		
prukr	5.5		
treng	5.9	trsp eng	5.3
trukr	6.8	trsp ukr	6.2
		ivsp eng	3.6
		ivsp ukr	5.3

Таблиця 4.7

#### ТЕМП ЧИТАНОГО ТА СПОНТАННОГО МОВЛЕННЯ

На основі отриманих даних (див. табл. 4.7) можемо спостерігати, що в усіх дев'ятох дикторів показники темпу вищі при мовленні рідною мовою (різниця для читаного мовлення – 0.9-2.0 скл/с, для спонтанного – 0.8-2.2 скл/с). При цьому в середньому для спонтанного мовлення різниця темпу для англійської й української мови була більшою, ніж для читаного мовлення.

Враховуючи інформацію про дослідження, які свідчать про нижчий темп мовлення в українській мові, протилежні показники в аналізованих нами дикторів, а саме – швидший темп мовлення рідною мовою, імовірно означають кращий рівень володіння дикторів цією мовою. Це збігається зі спостереженнями й висновками, які ми отримали при аналізі пауз: у мовленні рідною мовою в більшості дикторів тривалість пауз була меншою.

Різниця в темпі мовлення, як і в випадку з паузами, імовірно спричинена тим, що певна частина свідомої уваги дикторів виділяється на те, аби слідкувати за вимовою слів, коли в рідній мові такого не відбувається, адже в цій мові контроль вимови звуків відбувається на підсвідомому рівні. Отже, таке розподілення уваги на різні аспекти мовлення вповільнює синтез мовлення

дикторами загалом, що впливає в тому числі й на темп мовлення: у другій мові він уповільнюється.

#### **Висновки до розділу 4**

У цьому розділі ми проаналізували й порівняли деякі просодичні показники англійського й українського мовлення. Аналіз показав, що в залежності від мови частота основного тону не змінюється. Спостерігаємо незначні зміни: 3-16 Гц для читаного мовлення й 4-7 Гц для спонтанного мовлення. Але зміни не відбуваються в одному напрямку, тобто для певних дикторів ЧОТ зростає в англійському мовленні, в інших навпаки – в українському. Це не дозволяє нам виділити закономірності, які спричинені саме зміною мови на противагу індивідуальним характеристикам мовців.

Щодо довжини синтагм, то єдина закономірність, яку можемо спостерігати в читаному мовленні, – збільшення в більшості дикторів в англійському мовленні кількості коротких синтагм на 4-23% і зменшення кількості середніх на 1-18%. Але ця закономірність швидше за все спричинена відмінностями в особливостях синтаксичних структур двох аналізованих мов. Для спонтанного мовлення результати майже протилежні. В англійському мовленні у 5 з 6 дикторів спостерігаємо збільшення кількості довгих синтагм на 4-20% і зменшення кількості коротких синтагм на 10-34%. Враховуючи інформацію про збільшення тривалості пауз можемо припустити, що такі зміни викликані бажанням дикторів ретельніше продумувати свої фрази в англійському мовленні перед тим, як їх виголошувати.

Інші показники – інтенсивність, тривалість пауз, темп – змінюються закономірно для всіх або більшості дикторів.

Середній показник інтенсивності в читаному мовленні в п'яти з восьми дикторів вищий в українській мові на 1-5 дБ, у спонтанному мовленні – у п'яти з шести дикторів вищий на 3-7 дБ. Можемо говорити про закономірність такої зміни, оскільки вона актуальна для більшості дикторів.

Щодо темпу мовлення, то в усіх дев'ятьох дикторів показники вищі при мовленні рідною мовою. Для читаного мовлення різниця варіює в межах 0.9 скл/с і 2.0 скл/с, для спонтанного – 0.8 скл/с і 2.2 скл/с. При цьому в середньому для спонтанного мовлення різниця темпу в англійському й українському мовленні більша, ніж для читаного мовлення.

Перехід дикторів на англійську мову під час читаного й спонтанного мовлення завжди супроводжується збільшенням тривалості пауз. У більшості дикторів кількість довгих пауз збільшується (на 4-17%) за рахунок значного зменшення кількості середніх пауз (на 3-11%) і незначного зменшення кількості коротких пауз (на 1-7%). У деяких дикторів збільшується кількість середніх пауз (на 8%) за рахунок зменшення кількості коротких пауз (на 7%). У спонтанному мовленні тривалість пауз збільшується в більшій мірі, ніж у читаному: кількість довгих пауз збільшується на 2-24%, кількість середніх зменшується на 5-19%.

## ВИСНОВКИ

Для виявлення спільних і відмінних рис мовлення білінгвів у різних мовах ми дослідили такі просодичні характеристики, як паузація, довжина синтагм, частота основного тону й інтенсивність, темп мовлення. У якості методів аналізу ми використовували можливості мови скриптів *Praat* і мови програмування *Python* як у процесі створення анотацій до звукових записів, так і для подання показників обчислень у більш наочній і зручній для аналізу формі.

У результаті аналізу ми зробили певні висновки.

Складнощі продукування мовлення дикторами, спричинені різницею в рівні володіння другою (іноземною) мовою, імовірно виражаються в таких просодичних характеристиках, як паузація, темп та ІЗС. У нерідній мові в більшості дикторів зростає тривалість пауз і в усіх дикторів уповільнюється темп мовлення. Припускаємо, що це виникає у зв'язку зі збільшенням часу, потрібного дикторам на продукування англійського мовлення: підбір правильних слів, формування синтаксичних конструкцій, слідкування за вимовою (для спонтанного мовлення). У читаному мовленні цих процесів не відбувається, окрім останнього, який все одно спричиняє подовження пауз та уповільнення мовлення, але в меншій мірі, ніж у спонтанному мовленні.

У близько 70% випадків (10 із 14), ІЗС дикторів зростає при мовленні першою мовою (як читаному, так і спонтанному). Припускаємо, що це зумовлено тим, що диктор почуває себе вільніше й більш обізнаним у знанні рідної мови, тому в підвищеній інтенсивності голосу виражається більша впевненість у тому, щоб надавати мовленню більш яскраве емоційне насичення.

Що стосується показника ЧОТ, то він майже не залежить від мови, якою говорить диктор. Діапазон частоти основного тону для кожного мовця залишається тим самим, а зміна інтонаційного оформлення висловлень залежить від особливостей структури самого тексту, а також від особливостей

структури мови (незалежно від індивідуальних мовленнєвих характеристик диктора).

У відмінностях довжини синтагм дикторів для різних мов можна виділити певні закономірності, але вони виявляються лише у спонтанному мовленні: кількість довгих синтагм збільшується, а кількість коротких – зменшується, що можна пояснити більшою невпевненістю дикторів у своїй здібності виразити власні думки другою мовою та, як наслідок, потребою продумувати фрази більш детально перед їх виголошенням.

Серед перспектив дослідження цікавим є подальший аналіз українського та англійського мовлення українців-білінгвів в аспекті їх розподілу за гендерними ознаками. Також окрім просодичних показників, які вже бралися до уваги в цій роботі, можна досліджувати інтонаційні контури різних типів синтагм, паузи хезитації та звуки, які в них використовуються (е-кання, и-кання, ме-кання, а-кання, ам-кання, розтягування слів у якості хезитації), а також екстралінгвістичні параметри (кашель, сміх, прочищення горла і т.п.) в усному мовленні білінгвів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ануфрієва Н.Д. Мовна інтерференція в юридичному підстилі українського ділового мовлення: лексико-граматичний аспект // Одеський лінгвістичний вісник, 2014. № 3. С. 19-24.
2. Ахунзянов Э.М. Морфологическая интерференция в условиях русско-татарского двуязычия. Москва, 1976. 145 с.
3. Бабич В. Зіставлення мов як спосіб їх вивчення. Фонетика // Освітня політика: портал громадських експертів: [Електронний ресурс]. 2019. URL: <http://education-ua.org/ru/blogs/tema-1/1360-zistavlennya-mov-yak-sposib-yik-h-vyvchennya-fonetyka> (дата звернення: 20.05.2021).
4. Блягоз З.У. Проблема заимствования в адыгейском литературном языке // Филологический вестник АГУ. 2001. (№2). С. 78-81.
5. Валігура О.Р. Лінгвокогнітивні і комунікативні основи фонетичної інтерференції (експериментально-фонетичне дослідження англійського мовлення українців): автореф. дис. докт. філол. наук: спец. 10.02.04 "Германські мови", спец. 10.02.15 "Загальне мовознавство"/ О. Р. Валігура. Київ, 2010. 32 с.
6. Валігура О.Р. Фонетична інтерференція в англійському мовленні українських білінгвів: Монографія / О.Р. Валігура. Тернопіль: Підручники і посібники, 2008. 288 с.
7. Вежбицка А. Язык. Культура. Познание / за ред. Кронгауз М. А. Москва: Русские словари, 1996. 416 с.
8. Вивчаємо теоретичну фонетику: конспекти лекцій / Малімонова Н.В., Меркулова Т.К. Харків: ХНУ імені В. Н. Каразіна, 2004. 84 с.
9. Дешериев Ю.Д. Основные аспекты исследования двуязычия и многоязычия // Проблемы двуязычия и многоязычия: сб. науч. тр. Москва, 1972. С. 28-29.
10. Жлуктенко Ю.А. Лингвистические аспекты двуязычия. Киев: Вища школа, 1974. 164 с.
11. Маслова В.А. Введение в лингвокультурологию. Москва: Издательский центр «Академия», 2001. 208 с.

- 12.Мацнєва Є.А. Проблема дефініції терміна «білінгвізм» // Інтелект. Особистість. Цивілізація. 2016. № 12. С. 36-44.
- 13.Миличева М.Г. О психолінгвистике и нейролінгвистике билингвизма // Вопросы науки и образования, 2018. № 1. С. 173-174.
- 14.Общие сведения об ассоциациях и ассоциативных нормах: словарь ассоциативных норм русского языка / Леонтьев А.А. и др. ; под ред. Леонтьева А.А. Москва: Русский язык, 1977. 192 с.
- 15.Фрумкина Р.М. Психолінгвистика: учеб. пособие для студ. высш. учеб. заведений. Москва: Издательский центр «Академия», 2003. С. 165.
- 16.Холмогоров А.И. Конкретно-социологическое исследование двуязычия // Проблемы двуязычия и многоязычия: сб. науч. тр. Москва, 1972. С. 6.
- 17.Amenorvi C.R. Bilingualism and interference: a case study of ellipsis among some coordinate bilinguals // African journal of applied research, 2015. No. 2 Vol. 2. P. 117-125.
- 18.Ardila A. Bilingualism: A neglected and chaotic area, Aphasiology, 12:2. Miami, Florida, USA: Miami Institute of Psychology, 1998. Pp. 131-137.
- 19.Auer J. Bilingual conversation. Amsterdam and Philadelphia: John Benjamins, 1984. p. 116.
- 20.Bain B. Toward an integration of Piaget and Vygotsky: Bilingual considerations // Linguistics, 1975. No. 160 Vol. 13. P. 5-20.
- 21.Bloomfield L. Language. New York: Henry Holt, 1933. P. 56.
- 22.Cenoz J. Defining multilingualism // Annual Review of Applied Linguistics, 2013. № 33. P. 3-18.
- 23.Cummins J. Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters // Working Papers on Bilingualism, 1979. № 19. P. 21-129.
- 24.D'acierno M. R. Three Types of Bilingualism // Annual Meeting of the International Association of Teachers of English as a Foreign Language, March 27-30, Dublin, Ireland, 1990. 65 p.
- 25.Diebold R. Incipient bilingualism / In D. Hymes, (Ed.), Language in Culture and Society. New York: Harper & Row, 1964. P. 495-511.

26. Dzyubenko A.I., Bogatyreva M.A. On the phenomenon of bilingualism: from sociolinguistic and psycholinguistic perspective // *Baltic humanitarian journal*, 2020. No. 1 Vol. 9. P. 223-226.
27. Dwinastiti M. Language interference // Marlin Dwinastiti Wordpress: [Електронний ресурс]. 2013. URL: <https://marlindwinastiti.wordpress.com/2013/04/02/language-interference/> (дата звернення: 15.05.2021).
28. Grosjean F. *Studying Bilinguals*. Oxford: Oxford University Press, 2008. P. 14.
29. Hamers J., Blanc M. *Bilinguality and bilingualism*. Cambridge: Cambridge University Press, 2000. 458 p.
30. Haugen E. *The Norwegian language in America: A study in bilingual behavior*. Philadelphia: University of Pennsylvania Press, 1953. P. 7.
31. Kingdon R. *The Groundwork of English Intonation*. London, New York, Toronto: Longmans, Green and Co, 1958. P. 31-40.
32. Kochubei V. Ukrainian English Accent: Roots, Reasons and Basic Features // *Linguistische Treffen in Wrocław*, 2019. No. 2 Vol. 16. P. 271–279.
33. Landsberry L. Defining bilingualism, 2019. P. 145-154. URL : <https://core.ac.uk/download/pdf/268407604.pdf> (дата звернення: 10.05.2021)
34. Lenhardtova L. An Experimental Study of Phonological Interference in the English of Slovaks // *International Review of Applied Linguistics*, 1993. No. 4 Vol. 31. P. 291–308.
35. Li W. Research perspectives on bilingualism and multilingualism // *The Blackwell handbook of research methods on bilingualism and multilingualism*. Oxford, UK: Blackwell, 2008. P. 3-17.
36. Mackey F.W. The description of bilingualism / In J. Fishman (Ed.), *Readings in the sociology of language*. The Hague: Mouton, 1962. P. 52.
37. Macnamara J. The bilingual's linguistic performance: A psychological overview, 1967 // *Journal of Social Issues*, Issues 23:2(58), 2010. P. 58-77.
38. Martínez G. *Mexican Americans and Language: Del dicho al hecho*. Tucson, USA: The University of Arizona Press, 2006. 128 p.
39. Moradi, H. An Investigation through Different Types of Bilinguals and Bilingualism // *International Journal of Humanities & Social Science Studies*

- (IJHSSS), A Peer Reviewed Bimonthly Bilingual Research Journal. Chandigarh: Panjab University Press, 2014. No. 1 Vol. 2. P. 107-112.
40. Myers-Scotton C. Multiple Voices: An Introduction to Bilingualism. MA: Blackwell Publishing, 2009. P. 44.
41. Petrishin O. L., Luzhetska O. M. Features of Ukrainian-English prosodic interference // *Filologia, socjologia i kulturoznawstwo*, Warszawa: Diamond trading tour, 2016. P. 50-55.
42. Praat // Wikipedia: The Free Encyclopedia: [Электронный ресурс]. URL: <https://en.wikipedia.org/wiki/Praat> (дата зверення : 25.05.2021)
43. Romaine S. Bilingual language development / M. Barrett (Ed.) *The Development of Language*, Sussex, UK: Psychology Press, 1999. P. 251-275.
44. Sirbu A. Language interference triggered by bilingualism // "Mircea cel Batran" Naval Academy Scientific Bulletin, 2015. No. 1 Vol. 18. P. 374-376.
45. Sound: To Intensity..: [Электронный ресурс]. 2010. URL: [https://www.fon.hum.uva.nl/praat/manual/Sound\\_\\_To\\_Intensity\\_\\_.html](https://www.fon.hum.uva.nl/praat/manual/Sound__To_Intensity__.html) (дата звернення: 22.04.2021)
46. Sound: To Pitch..: [Электронный ресурс]. 2003. URL: [https://www.fon.hum.uva.nl/praat/manual/Sound\\_\\_To\\_Pitch\\_\\_.html](https://www.fon.hum.uva.nl/praat/manual/Sound__To_Pitch__.html) (дата звернення: 21.04.2021)
47. Weinreich U. *Languages in contact: Findings and problems*. The Hague: Mouton, 1968. P. 1.

## ДОДАТКИ

### Додаток 1

#### Відомості про дикторів

##### 1. Троїцька Валерія

Вік	20 років
Освіта	студентка 4 курсу, спеціальність «прикладна лінгвістика та англійська мова»
Місце народження	с. Городище-1, Луцький район, Волинська обл.
Місце проживання	м. Київ
Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	C2

##### 2. Гришин Ілля

Вік	21 рік
Освіта	студент 4 курсу, спеціальність «прикладна лінгвістика та англійська мова»
Місце народження	м. Кропивницький
Місце проживання	м. Київ
Перша мова (рідна)	Українська, російська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	C1

##### 3. Приймак Оксана

Вік	36 років
Освіта	повна вища, к. філол. наук, викладає англ. мову
Місце народження	Хмельницька обл.
Місце проживання	м. Київ
Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	C1

#### 4. Мендела Ірена

Вік	45 років
Освіта	повна вища, вчитель англ. мови
Місце народження	м. Калуш, Івано-Франківська обл.
Місце проживання	м. Калуш, Івано-Франківська обл.
Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	C1

#### 5. Кузів Юлія

Вік	20 років
Освіта	студентка 4 курсу, спеціальність «прикладна лінгвістика та англійська мова»
Місце народження	с. Станькова, Калуський район, Івано-Франківська обл.
Місце проживання	м. Київ

Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	B2

#### 6. Іванишин Назар

Вік	19 років
Освіта	студент 3 курсу, спеціальність «міжнародні відносини»
Місце народження	м. Калуш, Івано-Франківська обл.
Місце проживання	м. Краків, Польща
Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	B1-B2

#### 7. Ананевич Владислав

Вік	18 років
Освіта	повна середня
Місце народження	м. Калуш, Івано-Франківська обл.
Місце проживання	м. Торонто, Канада
Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	B2

## 8. Копос Аліна

Вік	23 роки
Освіта	неповна вища, вчитель англ. мови та літератури
Місце народження	м. Прага, Чехія
Місце проживання	м. Дублін, Ірландія
Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	C2

## 9. Матійців Софія

Вік	23 роки
Освіта	повна вища, спеціальність «міжнародні відносини»
Місце народження	м. Калуш, Івано-Франківська обл.
Місце проживання	м. Калуш, Івано-Франківська обл.
Перша мова (рідна)	Українська
Друга мова (нерідна)	Англійська
Рівень володіння другою мовою (на думку самого диктора)	B2

### Список тем для спонтанного мовлення дикторів

Список тем для спонтанного мовлення дикторів знаходиться за посиланням:

<https://docs.google.com/document/d/1JrH7zH-BAHQVpRsLbykBiI7ZYzIC23i4OnqyXxpaCts/edit>

У файлі Topics.docx знаходяться 15 тем, які диктори за власним бажанням могли використати для запису усного мовлення. Усі теми представлені англійською мовою, до кожної із них додається невеликий список навідних питань.

**Скрипт boundaries\_lv3-4.praat**

(скрипт встановлює межі на рівнях 3 (pitch) і 4 (intensity) відповідно до меж на рівні 1 (phrase) для файла .TextGrid)

```
if numberOfSelected ("TextGrid") <> 1
    exitScript: "Please select a Sound and a TextGrid first."
endif

textgrid = selected ("TextGrid")
selectObject: textgrid
n = Get number of intervals: 1

for i to n
    if i <> 1
        phrase_start = Get start time of interval: 1, i
        Insert boundary: 3, phrase_start
        Insert boundary: 4, phrase_start
    endif
endfor
```

**Скрипт meanpitch.praat**

(скрипт обчислення середньої ЧОТ для кожної синтагми для файла .TextGrid)

```
if numberOfSelected ("Sound") <> 1 or numberOfSelected
("TextGrid") <> 1
    exitScript: "Please select a Sound and a TextGrid first."
endif
sound = selected ("Sound")
textgrid = selected ("TextGrid")
selectObject: sound
pitch = To Pitch: 0.0, 50, 700
selectObject: textgrid
n = Get number of intervals: 1
for i to n
    tekst$ = Get label of interval: 1, i
    if tekst$ <> ""
        t1 = Get starting point: 1, i
        t2 = Get end point: 1, i
        selectObject: pitch
        f0 = Get mean: t1, t2, "Hertz"
        selectObject: textgrid
        Set interval text: 3, i, string$ (round (f0))
    endif
endif
endfor
```

**Скрипт meanintensity.praat**

(скрипт обчислення середньої інтенсивності для кожної синтагми для файла  
.TextGrid)

```
if numberOfSelected ("Sound") <> 1 or numberOfSelected  
("TextGrid") <> 1  
    exitScript: "Please select a Sound and a TextGrid first."  
endif  
sound = selected ("Sound")  
textgrid = selected ("TextGrid")  
selectObject: sound  
intensity = To Intensity: 100, 0.001  
selectObject: textgrid  
n = Get number of intervals: 1  
for i to n  
    tekst$ = Get label of interval: 1, i  
    if tekst$ <> ""  
        t1 = Get starting point: 1, i  
        t2 = Get end point: 1, i  
        selectObject: intensity  
        tns = Get mean: t1, t2, "dB"  
        selectObject: textgrid  
        Set interval text: 4, i, string$ (round (tns))  
    endif  
endfor
```

## Скрипт mean\_all.praat

(скрипт обчислення загальної середньої ЧОТ та інтенсивності для файла

.TextGrid)

```

if numberOfSelected ("TextGrid") <> 1
    exitScript: "Please select a TextGrid first."
endif

textgrid = selected ("TextGrid")
selectObject: textgrid

pitch_sum = 0
intensity_sum = 0
phrases = 0

n = Get number of intervals: 1

for i to n

    label$ = Get label of interval: 1, i
    if label$ <> ""

        pitch$ = Get label of interval: 3, i
        intensity$ = Get label of interval: 4, i
        pitch = number (pitch$)
        intensity = number (intensity$)
        pitch_sum += pitch
        intensity_sum += intensity
        phrases += 1

    endif
endfor

mean_pitch = pitch_sum / phrases
mean_intensity = intensity_sum / phrases

writeInfoLine: "Середня ЧОТ: " + string$ (mean_pitch)
appendInfoLine: "Середня інтенсивність: " + string$
(mean_intensity)
appendInfoLine: string$ (pitch_sum) + ", " + string$
(intensity_sum) + ", " + string$ (phrases)

```

## Скрипт count\_pauses.praat

(скрипт обчислення кількості пауз та поділу за їх тривалістю для файла

.TextGrid)

```

if numberOfSelected ("TextGrid") <> 1
    exitScript: "Please select a TextGrid first."
endif

textgrid = selected ("TextGrid")
selectObject: textgrid

short = 0
med = 0
long = 0

n = Get number of intervals: 1

for i to n

    label$ = Get label of interval: 1, i
    if label$ = ""

        ts = Get start point: 1, i
        te = Get end point: 1, i
        pause_length = te - ts

        if pause_length < 0.150
            short += 1
        elsif pause_length > 0.3
            long += 1
        else
            med += 1
        endif
    endif
endfor

total = short + med + long

objName$ = selected$: "TextGrid"
appendFileLine: "spontan_pauses.csv", objName$ + "," + string$
(short) + "," + string$ (med) + "," + string$ (long) + "," +
string$ (total) + "," + string$ (round (short/total*100)) + "%," +
string$ (round (med/total*100)) + "%," + string$ (round
(long/total*100)) + "%"

```

**Скрипт get\_phrases.praat**

(програма обчислення кількості пауз та поділу за їх тривалістю для файла  
.TextGrid)

```
if numberOfSelected ("TextGrid") <> 1
    exitScript: "Please select a TextGrid first."
endif

textgrid = selected ("TextGrid")
selectObject: textgrid

n = Get number of intervals: 1
objName$ = selected$: "TextGrid"

for i to n

    label$ = Get label of interval: 1, i
    if label$ <> ""

        appendFileLine: "./phrases_txt/" + objName$ + ".txt",
label$

    endif
endfor
```

## Програма edit\_phrases.py

(програма поділу синтагм на фонетичні слова)

```

import re, os

path = r'.\phrases_txt'
new_path = r'.\phrases_edited'
dir = os.listdir(path)

regex = {
    True: [r"(\b[''\w-]{1,4}\b) (\b[''\w-]+)", r'\1\2'], #
    False: [r"(\b[''\w-]{1,3}\b) (\b[''\w-]+)", r'\1\2'] #
}

for filename in dir:
    # with open(path + '\\ ' + filename, 'r', encoding='utf-16-be')
    as file:
        # phrases = file.read()
        #
        # with open(path + '\\ ' + filename, 'w', encoding='utf-8') as
        file:
            # file.write(phrases)

    with open(path + '\\ ' + filename, 'r', encoding='utf-8') as
    file:
        phrases = file.readlines()

    for n in range(len(phrases)):
        phrases[n] = phrases[n].strip('\n')
        if phrases[n] != '':
            phrases[n] = re.sub(regex['eng' in filename][0],
                                regex['eng' in filename][1],
                                phrases[n])

    with open(new_path + '\\ ' + filename, 'w', encoding='utf-8')
    as file:
        file.write('\n'.join(phrases))

```

### Програма `count_phrases.py`

(програма обчислення кількості синтагм за довжиною та збереження отриманих даних в таблиці `.csv`)

Код програми знаходиться за посиланням:

<https://drive.google.com/file/d/1bim37sDYPjp8TszvnT-3-BnSwSMmaqE1/view?usp=sharing>

Програма обчислює кількість кожного типу синтагм (однослівні, двослівні і т.д.) у текстовому файлі й зберігає інформацію у файлах формату `.csv`.

## Кількість синтагм за довжиною

(таблиці .csv з даними щодо довжини синтагм за довжиною)

Створені таблиці знаходяться за посиланням:

[https://drive.google.com/drive/folders/1d4sULToZPq\\_5qrVdPLRK11WrrRUJ](https://drive.google.com/drive/folders/1d4sULToZPq_5qrVdPLRK11WrrRUJ)

[O\\_Zd?usp=sharing](#)

У папці Таблиці.csv можна знайти два файли: *phrases\_spontan.csv* (таблиця з отриманими даними аналізу синтагм читаного мовлення дикторів) та *phrases\_read.csv* (таблиця з отриманими даними аналізу спонтанного мовлення дикторів).

Кожна з таблиць має однакову структуру для читабельності отриманих даних: Таблиці мають таку структуру: *filename* (назва файлу *.TextGrid*), *one* (кількість однослівних синтагм), *two* (к-сть двослівних синтагм), *three* (к-сть трислівних синтагм), *four* (к-сть чотирислівних синтагм), *five* (к-сть п'ятислівних синтагм), *six* (к-сть шестислівних синтагм), *seven* (к-сть семислівних синтагм), *eight* (к-сть восьмислівних синтагм), *nine* (к-сть дев'ятислівних синтагм), *ten<=* (к-сть синтагм довжиною 10 і більше слів), *total* (загальна к-сть синтагм у файлі), *one%* (частка однослівних синтагм у відсотках від загальної к-сті), *two%* (частка двослівних синтагм), *three%* (частка трислівних синтагм), *four%* (частка сотирислівних синтагм), *five%* (частка п'ятислівних синтагм), *six%* (частка шестисліних синтагм), *seven%* (частка семислівних синтагм), *eight%* (частка восьмислівних синтагм), *nine%* (частка дев'ятислівних синтагм), *ten<=%* (частка синтагм довжиною 10 і більше слів).