

АНАЛІЗ ТЕМПОРАЛЬНИХ ПЕРЕТВОРЕНЬ МОВНИХ ЕЛЕМЕНТІВ ДЛЯ ЗАДАЧ ЧАСОВОГО МАСШТАБУВАННЯ ГОЛОСОВИХ ПОВІДОМЛЕНЬ

Зореслава Шпак, Юрій Рашкевич

Національний університет "Львівська політехніка"

вул. Ст. Бандери, 12, м. Львів-13, 79646; тел.: (38-0322) 398-793; факс: 744-143

Електронна пошта: zshpak@polynet.lviv.ua; rashkev@polynet.lviv.ua

ABSTRACT

The characteristics of the speech units (sounds, syllables, words, etc.) duration have been specified for different pronunciation types. Time-scale transformation regularity of the Ukrainian voice phonemes sounding at different speech rates has been investigated. Six groups of oral speech elements are determined: stressed vowels, unstressed vowels, voiced consonants, unvoiced consonants, plosive stops and word-spacing pauses. The elements of specified groups are characterized by similar temporal transformations. Non-linear analytic function presenting the relative temporal changes of sound duration have been plotted for every group.

1. ВСТУП

Підвищення ефективності усномовної комунікації в людино-машинних системах ставить задачу керування темпом надходження голосової інформації. Щоб забезпечити розбірливість і натуральність звучання мови, відтвореної зі зміненою швидкістю, у процесах часового масштабування голосових записів треба враховувати природні темпоральні властивості мовних елементів [1].

У роботі наведено часові параметри структурних одиниць мовного потоку для різних темпів мовлення. Проаналізовано закономірності темпоральних змін

звуків української мови і виділено класи мовних елементів, що характеризуються спільністю перетворень, викликаних зміною швидкості мовлення. Побудовано графіки, що відображають залежність відносної зміни тривалості елементів виділених класів від загального коефіцієнта зміни темпу.

2. ЧАСОВІ ХАРАКТЕРИСТИКИ РІЗНИХ ТЕМПІВ МОВЛЕННЯ

Для визначення темпоральних параметрів мовлення використано набір текстів, які відображають сучасні тенденції у структурі української розмовної, ділової та наукової мови. Кожен з текстів (загальний обсяг мовного матеріалу – 167 речень та відповідно: 2138 слів, 5331 складів і 12315 фонем) промовлявся групою з семи мовців жіночої та чоловічої статі в чотирьох темпах: швидкому, звичайному, повільному та протяжному. У табл.1 наведено темпоральні дані двох мовців (M1 і M2), які відрізняються стилями мовлення та усереднені значення для всієї групи мовців.

Основні підсумки отриманих результатів:

- найбільш стабільною і незалежною від тексту характеристикою швидкості мовлення є параметр "складів за секунду" – його доцільно застосовувати як базову темпоральну характеристику;

Таблиця 1

Часові характеристики основних стилів мовлення

Параметр		Темп мовлення								
		швидкий			звичайний			повільний		
		M1	M2	сер.	M1	M2	сер.	M1	M2	сер.
Середня тривалість	речення (с)	8,16	5,29	5,79	12,43	5,10	7,57	20,34	9,69	12,08
	синтагми (с)	2,37	1,47	1,68	3,53	2,34	2,47	5,64	2,44	3,45
	слова (мс)	552	341	415	734	441	540	1106	556	667
	складу (мс)	221	131	159	302	173	206	419	213	267
	звуку (мс)	94	56	68	129	74	88	179	91	114
Швидкість мовлення	слів/хвилину	96,5	135,8	130,1	64,6	119,7	89,1	39,2	86,1	62,3
	складів/секунду	4,1	6,7	5,3	2,6	4,7	3,7	1,6	3,4	2,5
	звуків/секунду	9,2	16,3	13,2	6,3	11,5	9,0	3,7	8,2	6,2
Коефіцієнт прискорення/сповільнення		1,51	1,43	1,46	1,0	1,0	1,0	1,64	1,39	1,43

- варіативність темпоральних показників мовців є достатньо широкою, особливо у випадках швидкої та повільної вимови. Так, М1 володіє достатньо широким діапазоном зміни темпу і помірною швидкістю мовлення, а М2 може бути віднесений до групи швидких мовців – властивий для нього звичайний темп є близьким до швидкого мовлення М1;

- незважаючи на відмінності в абсолютних параметрах швидкості мовлення, всім мовцям властиві достатньо близькі значення відносної зміни темпу. Середні по всій групі коефіцієнти максимального прискорення та сповільнення склали відповідно 1,46 та 1,48 (коефіцієнти сповільнення характеризувались значно ширшим розкидом значень, ніж коефіцієнти прискорення);

- часові параметри мовця не залежать від його статі, а визначаються властивим даній людині стилем мовлення. Усім мовцям, які не мали спеціальної дикторської практики, важко було довготривало відхилятися від звичайної швидкості вимови – чим довшим був текст, тим меншим ставав загальний коефіцієнт зміни темпу.

3. ТЕМПОРАЛЬНІ ПЕРЕТВОРЕННЯ ПАУЗ І ЗВУКІВ

3.1. Темпоральні перетворення пауз.

Зміна темпу мовлення пов'язана з перерозподілом структури мовного потоку (рис.1). Перехід від звичайної мови до швидкої чи повільної насамперед відображається в зміні кількості та тривалості пауз: кількість пауз зменшується/зростає усереднено в 1,5 разів, а їх сумарна тривалість змінюється відповідно приблизно в 2,3 рази (табл.2). Водночас загальна тривалість звукової частини мовного потоку скорочується/збільшується тільки в 1,3 рази. Найбільший відсоток пауз у повільній мові – біля 30%. Відзначено, що при подальшому сповільненні темпу дещо зростає кількість пауз, але їх середня тривалість майже не змінюється. Як результат – у протяжній мові частка пауз зменшується до 20%.

Довгі синтагматичні паузи загалом характеризуються меншими і стабільнішими темпоральними змінами, ніж короткі міжсловні, які вирізняються найвищим коефіцієнтом варіації та є найбільш



Рис.1. Відсоток пауз і звуків у загальній тривалості мови

залежними від просодичних особливостей мовців. Середнє значення тривалості звукової ділянки між паузами, яке для експериментальних даних було близьким до 0,96 с, зазнає найменших темпоральних змін. Тільки в протяжному, близькому до співочого мовленні цей параметр зростає в 1,5 рази.

3.2. Зміни голосних звуків.

Для аналізу темпоральних перетворень звуків виконано аудіо-візуальний поділ мовних сигналів кожного із записів. Границі звуків контролювались та уточнювались за динамічними спектрограмами, що дозволило виділити складові частини звуків: початкові та кінцеві переходи і серединні квазі-стаціонарні ділянки.

Основна різниця відзначена в темпоральній поведінці голосних і приголосних (передусім шумних) звуків. Сповільнення темпу завжди проявлялось у зростанні частки голосних звуків – співвідношення голосні:приголосні для швидкого темпу складало 0,42:0,58; для звичайного – 0,45:0,55; для повільного – 0,51:0,49; а для протяжного – 0,59:0,41. У випадках переходу до швидкого чи повільного мовлення найбільше змінювались тривалості наголошених голосних (табл. 3), але в протяжній вимові спостерігалось невелике зменшення інтонаційного співвідношення між наголошеними і ненаголошеними звуками (рис.2). Місце голосного звуку в слові та

Темпоральні характеристики роздільчих пауз

Таблиця 2

Характеристика	Темп мовлення			
	швидкий	звичайний	повільний	протяжний
Тривалість пауз у 1с мовного потоку (мс)	95	198	293	214
Кількість пауз на слово	0,41	0,56	0,81	0,97
Середня тривалість паузи (мс)	141,6	238,1	383,4	405,7
СКВ тривалості пауз (мс)	129,1	188,7	215,3	273,7
Середня тривалість звукової ділянки між паузами (мс)	1012	664	905	1420
Відносне збільшення загальної тривалості пауз	1,0	2,47	5,73	6,41
Відносне збільшення середньої тривалості паузи	1,0	1,67	2,72	2,84

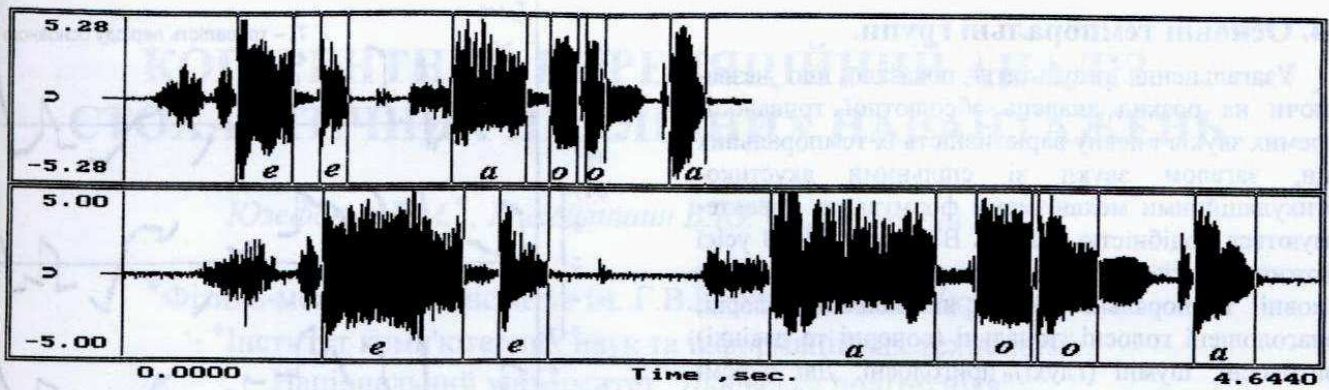


Рис.2. Мовні сигнали словосполучення *щобет жайворонка*, вимовленого в звичайному та протяжному темпах

просодика цього слова мали певний вплив на абсолютну тривалість звуку, проте відносні зміни тривалостей (ВЗТ) всіх голосних, викликані зміною темпу мовлення, були загалом достатньо однорідними.

3.3. Зміни приголосних звуків.

Відмінність у темпоральних перетвореннях голосних і приголосних звуків набуває принципового характеру у випадках значного сповільнення темпу (рис.2). Так, перехід від швидкого до протяжного мовлення пов'язаний із загальним збільшенням тривалості голосних у 4,4 рази, а приголосних – тільки в 2,2 рази. Водночас темпоральна поведінка різних приголосних звуків є теж достатньо неоднорідною (табл. 3). Проведено аналіз темпоральних змін для кожного звуку зокрема та узагальнено для основних класифікаційних груп приголосних [2]. В межах кожної групи враховувався вплив факторів пом'якшення, дзвінкості, назалізації та інших. Зроблено такі основні висновки:

- у випадках невеликих змін темпу (коефіцієнт прискорення/сповільнення не перевищує 1,4) всі приголосні змінюють свою тривалість приблизно однаково; при більших змінах темпу проявляються

відмінності в перетвореннях приголосних різної артикуляційної природи;

- перетворення сонорних звуків при невеликих змінах темпу є близькими до перетворень ненаголошених голосних, а при значних змінах темпу вони є ближчими до перетворень дзвінких приголосних;

- темпоральні зміни зімкнених звуків проявляються в неоднаковій зміні їх складових: паузи зімкнення та наступної короткої звукової ділянки (табл. 3);

- як для сонорних, так і для щільних та зімкнених звуків не встановлено істотних відмінностей у темпоральних змінах тривалостей м'яких та відповідних твердих звуків, за винятком випадків, коли м'які приголосні виступали як подовжені;

- приголосні з тональною складовою (дзвінкі звуки) дещо відрізняються за параметром тривалості, а також за значеннями ВЗТ від відповідних глухих приголосних (табл. 3), особливо у випадках протяжного мовлення;

- збільшення тривалості початково довгих звуків (зокрема подовжених приголосних) і скорочення початково коротких (зокрема вибухових) відбувається в меншій мірі, ніж для інших звуків даної групи.

Таблиця 3

Характеристики темпоральної тривалості звуків базових класів

Класи звуків		Темп мовлення												
		швидкий			звичайний			повільний			протяжний			
		ПВ, %	СТ, мс	ВЗТ	ПВ, %	СТ, мс	ВЗТ	ПВ, %	СТ, мс	ВЗТ	ПВ, %	СТ, мс	ВЗТ	
Голосні	ненаголошені	22,6	59,5	1,32	22,8	78,8	1,0	25,0	105,8	1,34	29,6	239,5	3,04	
	наголошені	19,5	91,1	1,42	21,9	129,6	1,0	25,2	181,5	1,40	29,8	408,5	3,15	
Приголосні	сонорні		20,6	49,6	1,30	19,9	64,7	1,0	17,5	84,6	1,31	15,1	119,9	1,85
	щільні	дзвінкі	4,8	61,7	1,27	4,6	78,6	1,0	4,1	98,9	1,26	3,4	136,8	1,74
		глухі	8,9	99,4	1,26	8,7	125,6	1,0	7,7	146,5	1,17	5,3	168,1	1,34
	зімкнені	дзвінкі	4,8	36,5	1,32	4,7	48,0	1,0	4,0	60,9	1,27	3,2	79,5	1,66
		глухі	2,6	24,6	1,25	2,5	30,7	1,0	2,1	36,6	1,19	1,2	42,3	1,38
		змикання	10,2	71,0	1,27	10,6	90,1	1,0	10,8	109,1	1,21	8,3	146,4	1,62
	африкати		2,8	77,1	1,28	2,6	98,4	1,0	2,4	110,8	1,13	1,7	127,6	1,30
подовжені		1,7	108,7	1,31	1,7	142,4	1,0	1,6	178,3	1,25	1,3	216,8	1,52	

Примітка: ПВ - процентний вміст; СТ - середня тривалість; ВЗТ - відносна зміна (збільшення/зменшення) тривалості.

3.4. Основні темпоральні групи.

Узагальнення результатів показало, що незважаючи на розкид значень абсолютної тривалості окремих звуків і певну варіативність їх темпоральних змін, загалом звуки зі спільними акустико-артикуляційними механізмами формування характеризуються подібністю значень ВЗТ (табл. 3). З усієї множини звуків української мови виділено чотири основні темпоральні групи: наголошені голосні, ненаголошені голосні, тональні (сонорні та дзвінкі) приголосні, шумні (глухі) приголосні. Дві окремі групи складають паузи зімкнення та роздільчі (міжслівні) паузи, темпоральні модифікації яких є принципово різними. На рис.3 наведено графіки ВЗТ мовних елементів основних темпоральних груп відносно значень, властивих для швидкого темпу (k – загальний коефіцієнт сповільнення вимови).

Зміна швидкості мовлення викликає неоднакові перетворення стаціонарної і перехідних ділянок звуків. Якщо зміни тривалості стаціонарних є достатньо регулярними, то трансформації перехідних ділянок характеризуються великою варіативністю. Загалом змінюються тривалості обох переходів, але усереднено в 1,4 (а в разі значних сповільнень – в 1,7) рази менше, ніж стаціонарних частин звуків.

В процесі досліджень проаналізовано вплив темпу мовлення на значення тривалості періодів основного тону (ПОТ) звуків, а також на форму мелодичного контуру фраз. Встановлено, що середні значення тривалості ПОТ тональних звуків і мелодика мови є практично незалежними від швидкості мовлення (рис.4). Зафіксовано тільки невелике зростання (біля 4%) середньої тривалості ПОТ у протяжному темпі та модуляції значень ПОТ для деяких довгих голосних. Темпорально незалежними виявилися також розподіли формантних частот і спектральні структури шумних звуків.

4. ВИСНОВКИ

Природна зміна швидкості мовлення пов'язана зі зміною тривалості всіх елементів мовного потоку. У разі невеликих змін темпу основних транс-

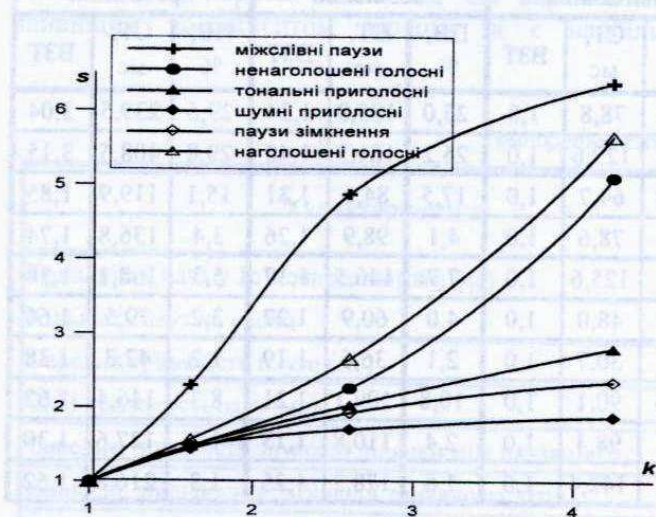


Рис.3. Графіки ВЗТ елементів основних темпоральних груп

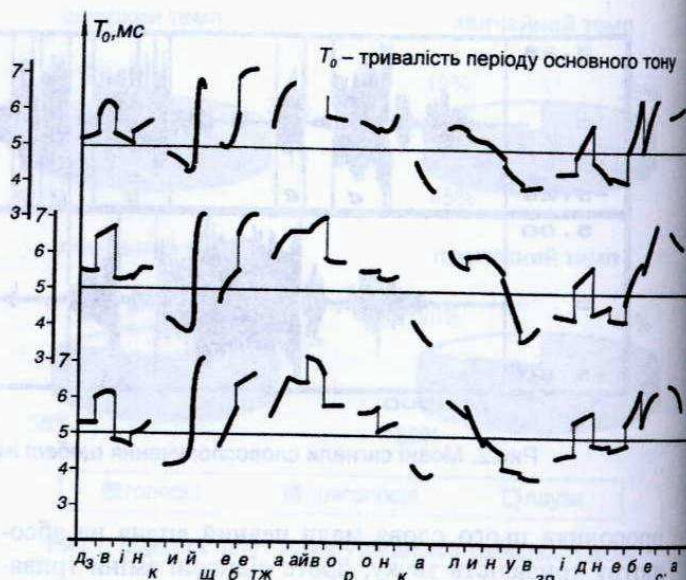


Рис.4. Мелодичні контури тестової фрази у звичайному (нижній рисунок), повільному (середній) і протяжному (верхній) темпах

формацій зазнають паузи, а модифікації всіх звуків є достатньо подібними і можуть бути описані лінійною залежністю від коефіцієнта зміни темпу. У випадках сповільнення/прискорення більше, ніж у 1,5 рази проявляються розбіжності у темпоральних перетвореннях звуків з різними механізмами формування. Ці відмінності стають принциповими для протяжної мови.

За близькістю значень ВЗТ і подібністю перетворення структур сигналів, викликаних зміною швидкості мовлення, всі звуки української мови можна об'єднати в чотири основні темпоральні групи; ще дві окремі групи формують паузи. Залежність ВЗТ елементів кожної з цих груп від величини загального коефіцієнта зміни темпу носить нелінійний характер. Особливості зміни тривалості надмірно довгих чи коротких звуків вимагають додаткового введення функції нормування [3]. Підтверджено, що основні спектральні характеристики звуків є темпорально стабільними.

Таким чином, реалізацію зміни швидкості відтворення мовних записів за умови натуральності звучання голосового повідомлення необхідно здійснювати через диференційовану зміну тривалості мовних елементів відповідно до властивих їм значень ВЗТ, зберігаючи загальну структуру звукових сигналів, передусім тональних.

ЛІТЕРАТУРА

1. Рашкевич Ю.М. *Перетворення часового масштабу мовних сигналів*. - Львів: Академічний експрес, 1997. - 140 с.
2. Тоцька Н.І. *Сучасна українська літературна мова*. - Київ: Вища школа, 1981. - 183 с.
3. З. Шпак. *Нормування тривалостей звуків у процесі сповільненого відтворення мовних повідомлень* // Вісник Націон. ун-ту "Львівська політехніка": Львів, 2001. - №433. - С. 266-271.