

Мовознавчий вісник: Зб. наук. праць на пошану професора Катерини Городенської з нагоди її 60-річчя / МОН України. Черкаський нац. ун-т ім. Б. Хмельницького; Відп. ред. Г.І. Мартинова. – Черкаси, 2009. – Вип. 8. – С. 167-176.

Лариса Азарова

УДК 004.43 (075.8)

Комп'ютерні технології у дослідженні гармонійності побудови складних слів

У статті обґрунтовано доцільність використання сучасних комп'ютерних технологій у рамках створення теорії гармонійної побудови складних слів в українській мові на засадах методології “золотої пропорції” та використання лінгво-інформаційного підходу; визначено основні принципи побудови складних слів на структурному й фонетичному рівнях.

Ключові слова: сучасні комп'ютерні технології, лінгво-інформаційний підхід, структурний та фонетичний рівень побудови складних слів, “золота пропорція”, композит, юкстапозит, гармонійні співвідношення.

У теперішній час мова значно розширює свої комунікативно-суспільні функції. Продовжуючи залишатися найважливішим знаряддям комунікації людей, природна мова стає тією основою, на базі якої здійснюється спілкування між людьми та їх групами з комп'ютером і комп'ютерними системами.

Бурхливий розвиток комп'ютерної техніки в останні десятиліття активно впливає на інформатизацію суспільства і сприяє поширенню використання комп'ютерних технологій у багатьох галузях людських знань, зокрема у мовознавстві. Основними напрямками використання комп'ютерних технологій у цій галузі є автоматизоване розпізнавання сенсу тексту та синтезу мовлення; комп'ютерне наближення слововживань тексту до їхньої канонічної форми; створення автоматизованих систем інформаційного пошуку; розроблення методів автоматичного анотування, комп'ютерне реферування та переклад тексту шляхом його компресії; лексичне забезпечення автоматизованих систем управління. Проблемам комп'ютерної лінгвістики, одного з новітніх напрямів інформаційної технології сучасного мовознавства, присвячені публікації Агеєва В. М., Узилевського Г. Я., Андрющенка В. М., Баранова А. М., Волошина В. Г. та ін.[1, 4 – 6]. Водночас ще не створено довершених універсальних теорій, які можна було б покласти в основу побудови практичних комп'ютерних програм. Тому розвиток нових комп'ютерних технологій у лінгвістиці є конче актуальним і необхідним. Концептуально дослідження пов'язане з науковою програмою “Закономірності розвитку мов і практика мовної діяльності”.

Метою статті є обґрунтування доцільності використання комп'ютерних технологій у рамках створення теорії гармонійної побудови складних слів в українській мові на засадах методології “золотої пропорції” та використання

лінгво-інформаційного підходу. Для досягнення цієї мети розв'язано такі завдання:

- розроблено основні положення теорії гармонійної побудови складних слів на засадах концепції “золотої пропорції”;

- запропоновано методологію дослідження гармонійної побудови складних слів, що спирається на лінгво-інформаційний підхід та сучасні комп'ютерні технології;

- визначено основні принципи побудови складних слів на структурному й фонетичному рівнях;

- висунуто гіпотезу про зв'язок виявлених гармонійних співвідношень з механізмом роботи мозку людини.

пропорціями побудови досліджуваних складних номінацій як на структурному, так і фонетичному рівнях;

- дослідження закономірностей побудови композитів і юкстапозитів у сучасній українській мові, а також висунення гіпотези про зв'язок виявлених гармонійних співвідношень з механізмом роботи мозку людини.

Для розв'язання вказаних завдань автор пропонує як базову форму дослідження словотвору використовувати комплексний лінгво-інформаційний підхід, який базується на використанні інформаційних комп'ютерних технологій. Суть його полягає в тому, що під час досліджень й аналізу складних слів використовуються не тільки традиційні лінгвістичні методи і прийоми, а й прийоми та методи теорії штучного інтелекту, теорії інформації комп'ютеризованої обробки лінгвістичної інформації тощо. Лінгво-інформаційний підхід ґрунтується на поєднанні як традиційних лінгвістичних методів і прийомів словотвору, так і запозичених методів і прийомів інформаційних технологій [2, 67].

Одним із таких комплексних підходів щодо розглянутої проблеми може слугувати методологія дослідження гармонії у лінгвістиці, зокрема у складних двокомпонентних композитах та юкстапозитах.

У методологічному аспекті для вимірювання довжини складного слова та його компонентів можна використовувати два підходи. Суть першого полягає в тому, що за одиницю вимірювання довжини беруть безпосередньо склад. При цьому довжина слова або його компонентів дорівнюватиме кількості складів. Цей підхід є зручним для аналізу номінативних одиниць, які писемно представлені в текстах.

Другий підхід полягає в тому, що за одиницю вимірювання довжини вибирають одиницю фізичної величини – часу, наприклад, мілісекунду. Це означає, що на фонетичному рівні номінативну одиницю розглядають у динаміці як відрізок потоку мовлення, яка акустично характеризується певною тривалістю. При цьому слід зазначити, що тривалість звучання композита (юкстапозита), як правило, не має сталого значення. Така тривалість залежить від низки чинників, зокрема, від характеру і темпу мовлення, який завжди певною мірою є індивідуальним у різних осіб, від інтонації мовця тощо. Уникнути вказаних незручностей можна, застосовуючи для аналізу відносний показник-пропорцію.

У випадку структурного рівня, коли номінативні одиниці зафіксовані у статичній формі і будь-які зміни їхньої довжини принципово неможливі, значення пропорції обчислюється за допомогою формули:

$$P = \frac{\text{кількість складів у двокомпонентному слові}}{\text{кількість складів у великому компоненті}}$$

Для обчислення пропорцій побудови складних одиниць на фонетичному (акустичному) рівні треба чітко визначити центр їх поділу на компоненти у вигляді часової відмітки. Виміряні значення тривалості звучання цілого слова і великого компонента фіксуються. При цьому значення пропорції обчислюється з відношення:

$$P_{\phi} = \frac{\text{тривалість вимовляння двокомпонентного слова (мс)}}{\text{тривалість вимовляння великого компонента (мс)}}$$

Оскільки пропорція P_{ϕ} являє собою відносну величину, то такий суб'єктивний чинник, як неоднаковий темп мовлення в різних осіб не може істотно впливати на точність оцінювання цього показника. Водночас, такі індивідуальні особливості окремої особи, як неоднакове почуття "гармонійності", "зсув логічного наголосу", можуть призвести до певного розмивання значень P_{ϕ} для слова, яке вимовляють різні особи. Тому для зменшення похибки в оцінюванні пропорцій звучання композитів і юкстапозитів під час проведення практичних дослідів було залучено фахівців-фонетистів, професійних дикторів. Саме в них правильно "поставлено" дикцію, вони чітко вимовляють слова, передусім їхні закінчення, мають розвинене почуття "часової пропорційності".

Систематизована таблиця значень пропорцій P , отримана в ході досліджень на структурному рівні, має вигляд:

P	П						
	0	2/1	4/2	6/3	8/4	10/5	12/6
1	3/2	5/3	7/4	8/5	10/6	11/7	13/8
2	4/3	6/4	9/6	13/9			
3	5/4	7/5	10/7	14/10			
4	6/5	8/6	11/8				
5	7/6	9/7	12/9				
6	8/7	10/8	13/10				

У наведеній таблиці параметр p пов'язано з так званими p -числами Фібоначчі [9, 35] і визначає характер зв'язку кожного наступного члена ряду з попередніми. Так, наприклад, якщо $p=0$, то кожний наступний член подвоюється щодо попереднього. Коли ж $p=1$, кожний наступний член, починаючи з другого, дорівнює сумі двох попередніх. Якщо $p=2$, то кожний наступний член дорівнює сумі попереднього і того, що розташований через

один. У загальному випадку, коли $p=k$, кожний наступний член дорівнює сумі попереднього й того, який виступає після $k-1$.

Схематично пропорції побудови композитів на структурному рівні можна продемонструвати на прикладі композита:

$$\begin{array}{l} 6 : 3 \\ \text{авіатехніка (}\Pi=6/3\text{)} \end{array} \quad \begin{array}{l} 6 : 4 \\ \text{юкстапозита бійці-штурмовики (}\Pi=6/4\text{)}. \end{array}$$

$$\begin{array}{l} 3 | 3 \\ 2 | 4 \end{array}$$

Центр поділу в композитах позначаємо вертикальною лінією знизу, а ліворуч і праворуч від неї проставляємо кількість складів у компонентах. Значення пропорції записуємо зверху над складною номінативною одиницею або праворуч у дужках. Для юкстапозитів центром поділу слугує дефіс, який поділяє слово на компоненти

Систематизована таблиця значень пропорцій Π_ϕ , отриманих у ході досліджень на фонетичному (акустичному) рівні, має таку форму:

P	Π_ϕ	Π'_ϕ
0	2,00	2,00
1	1,50; 1,67; 1,75; 1,60; 1,57; 1,625	1,618
2	1,33; 1,50; 1,44; 1,46	1,465
3	1,25; 1,40; 1,43	1,38
4	1,20; 1,38	1,32
5	1,17; 1,29	1,285
6	1,14; 1,30	1,26

Слід відзначити, що, якщо в межах структурного підходу використовувалися цілочислові відношення типу: $3/2$; $5/3$; $8/5$ і т.п., то в межах фонетичного для вимірювання часових інтервалів потрібно послуговуватися переважно дробовими показниками. Наведені в таблиці значення Π_ϕ і Π'_ϕ враховують p -числа Фібоначчі, числа Люка, а також значення “золотих” p -пропорцій [8, с.38]. Як дробові показники слід також використовувати корені узагальненого рівняння “золотих” S -пропорцій [3, с.38], зокрема, такі:

S	2	3	4	5
Π_ϕ	1,618	1,84	1,93	1,97

Вимірювання тривалостей звучання композитів, юкстапозитів та їхніх компонентів, а також визначення центру поділу досліджуваної номінативної одиниці на компоненти здійснюється за допомогою спеціальної апаратури. До складу такої апаратури входять пристрої запису й відтворення звукових сигналів, принципи дії яких, у значній мірі, ґрунтуються на теорії цифрового оброблення мовних сигналів. Певні досягнення у цій галузі має Вінницький національний технічний університет, зокрема, відділ аналого-цифрових систем. Розроблена апаратура дає змогу досить точно (похибка не

перевищує 1%) дослідити форму звукового сигналу (інтонограму), який виникає під час вимовляння тієї чи іншої номінативної одиниці, визначити центр поділу слова на компоненти, виміряти тривалість самого слова і його компонентів, зафіксувати дані в пам'яті комп'ютера. Озвучені номінативні одиниці, перетворені в цифрову форму і представлені на відповідному носії (компакт-диски, дискети), можна використовувати для подальших досліджень. Для їх реалізації достатньо наявності сучасного персонального комп'ютера, обладнаного засобами мультимедія, зокрема, навушниками або звуковими колонками, а також спеціалізованого програмного забезпечення. У цьому плані досить зручним є пакет прикладних програм "Wave for Windows", який виявився досить зручним, оскільки він розрахований не лише на фахівців з комп'ютерної техніки чи звукотехніки, а й на пересічного користувача комп'ютерних технологій, зокрема, на фахівця-лінгвіста.

Використання комп'ютерних технологій на фонетичному рівні можна продемонструвати на прикладі дослідження слова

437 мс 706 мс

авто|транспорт : $P_{\phi} = (437 \text{ мс} + 706 \text{ мс}) : 706 \text{ мс} \approx 1,619$

1143 мс

У цьому випадку тривалість звучання всього слова становить 1142 мс, великого компонента – 706 мс, а пропорція P_{ϕ} дорівнює $\approx 1,619$, що дає підставу вважати побудову цієї одиниці гармонійною і такою, що належить до класичної "золотої" пропорції.

Варто зауважити також, що пропорції P_{ϕ} і P на структурному і фонетичному рівнях можуть відрізнятися. Так, наприклад, композит-іменник *водосховище* ($P_{\phi} = 1,33$) на фонетичному рівні входить до сім'ї "золотої" 2-про- порції, а на письмі він має пропорцію $P=5/3$ і належить до класичної "золотої" 1-пропорції. Юкстапозит *люкс-секунда* ($P_{\phi} = 1,66$) на фонетичному рівні кваліфікуємо як одиницю сім'ї "золотої" 1-пропорції, а на структурному – ($P=4/3$) він належить до "золотої" 2-пропорції.

Проте пропорції більшості складних номінативних одиниць і під час вимови, і на письмі збігаються, утворюючи сім'ю класичної "золотої" 1-пропорції, наприклад:

3:2

5:3

дроворуб ($P=3/2$; $P_{\phi}=1,72$), хата-читальня ($P=5/3$; $P_{\phi}=1,57$),

2 1

2 3

5:3

нижченазваний ($P=5/3$; $P_{\phi}=1,68$).

2 3

Підсумки дослідження масивів слів, що входять до комп'ютеризованого словника, дають підставу стверджувати про гармонійність "конструкцій" складних номінативних одиниць у сучасній українській мові, оскільки кількісні значення пропорцій їх побудови відповідають числовим даним, наведеним у вищевказаних таблицях. Виявлені гармонійні закономірності у свою чергу підтверджують гіпотезу про глибинний зв'язок словотвору з механізмом роботи людського мозоку людини [7, с.56].

Висновки. Виявлені за допомогою комп'ютерних технологій закономірності гармонійної побудови складних номінативних одиниць на структурному і фонетичному рівнях мають науково-теоретичне і прикладне значення не тільки для вказаної категорії складних слів, а й словотвору в цілому.

Отримані кількісні показники гармонійної побудови складних слів доцільно поширити на такий напрям прикладної лінгвістики, як синтез людського голосу за допомогою комп'ютерних технологій і створення робота-диктора, який буде "ідеально" вимовляти слова, чітко ставитиме логічний наголос у реченнях, виділятиме гармонійний центр тощо.

Література

1. Агеев В. Н., Узилевский Г. Я. Человеко-компьютерное взаимодействие, процессы, модели. – М.: Мир книги, 1995. – 352 с.
2. Азарова Л. Є. Структурна та фонетична побудова складних одиниць у концепції "золотої" пропорції. Монографія. – Вінниця: Віноблдрукар-ня, 2001. – 284 с.
3. Азаров О. Д. Основи теорії аналого-цифрового перетворення на основі надлишкових позиційних систем числення. Монографія. – Вінниця: УНІВЕРСУМ – Вінниця, 2004. – 260 с.
4. Андрющенко В. М. Проблемы вычислительной лингвистики и автома-тическая обработка текстов на естественном языке. – М.: МГУ, 1980. – 184 с.
5. Баранов А. Н. Введение в прикладную лингвистику. – М.: Эдиторная УРСС, 2001. – 358 с.
6. Волошин В.Г. Комп'ютерна лінгвістика: Навчальний посібник. – Суми: ВТД "Університетська книга", 2004. – 382 с.
7. Соколов А. А., Соколов Я. А. Математические закономерности электрических колебаний мозга. – М.: Наука, 1975. – 98 с.
8. Стахов А. П. Новая математика для живой природы. Гиперболические функции Фибоначчи и Люка. – Винница, 2003. – 264 с.

Azarova E. Larisa, Gorchinskaya V. Ludmila, Kuharchuk V. Galina.
Computer technology as perspective technique of word construction researching.

In the article the using of computer technology as perspective technique of word construction researching in particular of bi-structural complicated nominations building on structural and phonetic levels are considered.

It's shown the results of researching of word sets which are being in the computer vocabulary allow to assert, that it is harmony constructions under the

conditions of qualitative harmony indexes correspond to the recurrent correlations like “gold” proportions.

МОНОГРАФІЯ

Структурна та фонетична побудова складних одиниць у концепції золотої пропорції. – Вінниця: ВЕГА, 2001. – 284 с.