

АНОТАЦІЯ

Парк С. Фонемна структура словоформ у готській мові: моделі побудови. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття ступеня доктора філософії за спеціальністю 035 Філологія. – Національний педагогічний університет імені М. П. Драгоманова, Київ, 2020.

Зміст анотації

У дисертації розроблено новий квантитативно-дистрибутивний підхід до вивчення фонемної структури словоформ у готській мові та їх комбінаторних моделей, з позицій якого доповнено теоретичні положення системно-структурної організації мовних одиниць усіх рівнів і, передовсім, фонологічного.

Крізь призму структурної фонології у роботі розвинуто ідеї кінемної теорії, згідно з якими кінема визначена як найменша артикуляційна одиниця, з одного боку, і найменша перцептивна одиниця, з іншого, а фонема становить стандартний для певної мови блок кінем.

У комбінаториці кінем розрізняють два види: комбінаторику внутрішньофонемну і комбінаторику міжфонемну. Внутрішньофонемна комбінаторика як голосних, так і приголосних фонем є сукупністю кінем. Тимчасом міжфонемна комбінаторика визначає можливість послідовної реалізації фонем у морфемі та словоформі. Правила першої і другої є специфічними для кожної окремої мови, які розкривають механізми 1) сполучуваності, 2) обмеженої сполучуваності та 3) несполучуваності фонем.

Система фонем готської мови відображає закони як праіндоєвропейської мови, так і ті фактори, що вплинули на її дивергенцію і були покладені в основу спільногерманської мови. Цей тривалий процес відбувався на тлі змін граматичного типу мови: від кореневого ізолятивного через стадію аглютинації до флективного ще в дописемну епоху. Релікти ознак, притаманні першим двом з цих граматичних типів, знайшли підтвердження у готських манускриптах.

Аналіз готського текстового матеріалу свідчить про розпад флективного ладу і початок динаміки мовної системи до аналітичного типу. На жаль, еволюційна лінія готської мови обірвалася на початку VIII ст. н.е., і наявні писемні пам'ятки IV-VI ст., що є найдавнішими серед інших давньогерманських мов, доступні дослідникам в одному синхронному стані мови.

Постульований у дослідженні квантитативно-дистрибутивний підхід надав змогу схарактеризувати фонемну структуру словоформ готської мови та їх фонемної комбінаторики на основі таких критеріїв, як довжина фонем, комбінаторика двох класів фонем – голосних та приголосних – в ініціальній, медіальній та фінальній позиціях.

За результатами цих характеристик було укладено картотеку дослідження, що складається з 9443 словоформ. Текстова частота цих словоформ у Готській Біблії та менших за обсягом готських рукописах становить 67464 слововживання.

Ідентифіковано 18 груп словоформ залежно від їх довжини у фонемах: макросистема готської мови накладає обмеження на реалізацію словоформ, довжиною більше ніж дев'ятнадцять фонем. Кожна із 18 груп представлена різною кількістю словоформ і різною частотою їх вживання у готських манускриптах. Не знайдено жодної вісімнадцятифонемної словоформи.

Зафіксовано, що понад 97% словоформ (9165) готської мови мають довжину від трьох до дванадцяти фонем. П'ять груп словоформ, довжиною у п'ять, шість, сім, вісім і дев'ять фонем, мають незначні розбіжності у своїй кількості й є найчисельнішими: з усіх словоформ їхня кількість складає 6849, що становить 72,5% від загальної кількості (9443) усіх словоформ. Кількість словоформ, довжиною від однієї до двох та від тринадцяти до дев'ятнадцяти фонем, у десятки разів менша і налічує 278 словоформ, що становить 2,9% від усієї кількості словоформ.

Доведено, що система готської мови накладає обмеження на реалізацію одно- та двофонемних словоформ. Ще більше обмежень мають словоформи, довжиною від 10 до 19 фонем: зі збільшенням довжини словоформ їхня продуктивність зменшується і залежить від кількості фонем у словоформі.

Укладено інвентар із 656 комбінаторних моделей словоформ готської мови, що є канонічними формами (КФ) і представляють фонемну структуру словоформи у вигляді ланцюжка реалізацій голосних (V) та приголосних (C) фонем. Теоретично можлива кількість таких моделей у словоформах готської мови, довжиною від однієї до дев'ятнадцяти фонем, складає понад один мільйон КФ.

Порівняльний аналіз теоретично можливих й імплементованих канонічних форм показав, що готська мова використовує потенціал своєї системи фонем менше, ніж на один відсоток, а кількість реалізованих канонічних форм зменшується зі збільшенням довжини словоформи у фонемах. Так, у готських словоформах, довжиною від однієї до шести фонем, мовна система використовує 50% теоретично можливих словоформ в одно- і двофонемних комбінаціях, 62,5% у трифонемних, 68,8% у чотирифонемних, 59% у п'ятифонемних і 51,5% у шестифонемних. Відсоток теоретично можливих словоформ зі збільшенням їх довжини – від семи фонем до дев'ятнадцяти – зменшується: для семифонемних словоформ він становить 32%, для восьмифонемних – 22%, для дев'ятифонемних – 15%, для десятифонемних – 8%, для одинадцятифонемних – 4,7%, для дванадцятифонемних – 2,2%. Починаючи з тринадцятифонемних словоформ відсоток теоретично можливих фонемних комбінаторних типів стрімко знижується і становить 0,8% у тринадцятифонемних моделях, 0,27% – у чотирнадцятифонемних, 0,05% – у п'ятнадцятифонемних, 0,009% – у шістнадцятифонемних, 0,002 – у сімнадцятифонемних. При цьому вісімнадцятифонемні моделі фонемної будови словоформ у готській мові не реалізовані загалом. Відсоток реалізації дев'ятнадцятифонемної словоформи становить 0,0002%.

Підтверджено припущення про те, що реалізація комбінаторних моделей фонемної структури словоформ (КФ) готської мови залежить від довжини словоформ у фонемах (кількості фонем, що входять до структури словоформи), від морфемної структури словоформи, від реалізації ініціальних (початкових), медіальних (серединних) та фінальних (кінцевих) сполучень голосних

і приголосних фонем у межах словоформи, а також від конфігурації внутрішньофонемної комбінаторики кінем, що входять до внутрішньої структури фонем як стандартні, унікальні та притаманні лише конкретній мові блоки кінем.

Виявлено, що система фонем і система фонематичних опозицій готської мови не допускають сполучення більше трьох голосних і чотирьох приголосних фонем у вокалічних і консонантних групах. Сполучуваність фонем у словоформах готської мови на рівні первинних фонологічних одиниць регулюється і залежить від внутрішньофонемної конфігурації кінем у внутрішній структурі кожної фонемі, а також від міжфонемної комбінаторики, в першу чергу, модальних, потім – локальних кінем. Фонемна система готської мови не допускає переобтяження консонантних груп маркованими кінемами.

Роль кожної КФ у системі мови визначається її моделювальною силою, яка представлена частотою використання словоформ певної моделі. *Установлено*, що найбільшу моделювальну силу мають 24 КФ, довжиною від однієї до одинадцяти фонем, які утворюють ядро системи. Частота їх використання у готських манускриптах становить 50660 слововживань (понад 75% частоти з усіх проаналізованих словоформ). До ядра системи входять 10 канонічних форм, які моделюють 25440 слововживань, що функціонують і як односкладові незв'язані кореневі морфеми, і як словоформи. 47 канонічних форм становлять основну систему словоформ і моделюють 10152 слововживання.

Текстова частота словоформ, побудованих на моделях 71 КФ, становить 60812 слововживань і складає понад 90% від сумарної частоти словоформ. Решта 585 канонічних форм утворюють периферію і моделюють 6652 слововживання, що становить 9,86% від сумарної частоти усіх зареєстрованих словоформ.

Зафіксовано 23 КФ (із 656), фонемна структура яких не ускладнена послідовною реалізацією ні голосних, ні приголосних фонем і які моделюють 2329 словоформ (24,66% від усієї кількості зареєстрованих словоформ), частота їх використання складає 36287 слововживань, що становить 53,79% сумарної частоти всіх словоформ.

Аналіз ініціальних вокалічних груп показує жорсткі фонологічні обмеження, що знижують сполучуваність голосних фонем в ініціальній позиції, особливо у словах германського походження. Реалізація вокалічних груп у біблійних власних назвах є результатом транскрипції цих імен із давньогрецького оригіналу.

Реалізація фінальних вокалічних груп зумовлена морфологічною системою готської мови: іменною, прономінальною та дієслівною парадигмами. Фонемна комбінаторика голосних фонем у медіальних вокалічних групах залежить від морфологічних і морфосинтаксичних чинників.

Установлено 33 двофонемні ініціальні, 75 фінальних і 48 трифонемних фінальних консонантних груп і, відповідно, 9,14%, 20,78% та 0,7% реалізації від кількості теоретично можливого конструктивного потенціалу. Текстова частота слівформ, що містять ці консонантні групи в готських рукописах, становить 5038, 5571 та 1315 слововживань, які моделюються 13, 141 та 138 канонічними формами. Три фонемні: /h^w, z, j/ взагалі не беруть участі в комбінаториці двофонемних ініціальних груп та чотири фонемні /p, h^w, w, j/ не зареєстровані в комбінаториці трифонемних консонантних груп.

У формуванні чотирифонемних фінальних консонантних груп беруть участь лише вісім фонем. Комбінаторика приголосних фонем в ініціальних і фінальних консонантних групах підпорядкована принципу сонорної ієрархії зі зростанням в ініціальній позиції та спадом – у фінальній.

Уперше *доведено*, що система готської мови віддає перевагу прикритим закритим (CVC) і прикритим відкритим складам (CV), що підтверджується їх високою частотою. Неприкриті закриті склади (VC) характерні для службових частин мови (часток, прийменників і сполучників) і для займенників, які були успадковані з праіндоєвропейської та спільногерманської мов.

Канонічними типами фонемної структури індоєвропейського кореня були односкладові морфеми CVC та VC. Коренева морфема CVC разом із фонемою суфіксу ставала двоскладовою основою з фонемною структурою CV-CV. Неприкриті відкриті склади реалізовані в ініціальній позиції слівформ.

Частота таких словоформ є низькою, оскільки цей тип складу не був притаманний праіндоєвропейській мові, не був він характерним і для спільногерманської мови, а тому не отримав функціональної підтримки і в готській мові.

Послідовна реалізація голосних і приголосних фонем у неприкритих закритих складах і приголосних у прикритих відкритих складах не супроводжувалася ні артикуляційними (вимовними, антропофонічними), ні аудитивними (перцептивними) труднощами, оскільки фонемна комбінаторика в таких словоформах не мала обмежень.

Застосування квантитативно-дистрибутивного методу з елементами статистичного аналізу підтвердило достовірність отриманих наукових результатів, що становить 99,5%.

Запропонована методика може бути використана в дослідженнях з фонології германських, інших близько- та віддаленоспоріднених індоєвропейських мов, а також мов, що належать до різних граматичних типів.

Ключові слова: дистрибуція фонем, канонічна форма, кінема, комбінаторика, модель, словоформа, фонема, фонемна структура, частота.

ABSTRACT

Park S. Phonemic structure of word-forms in Gothic: construction patterns. – Qualifying Academic Paper. Manuscript.

Thesis for the degree of higher education of Doctor of Philosophy, Specialty 035 Philology. – National Pedagogical Dragomanov University, Kyiv, 2020.

Abstract content

The dissertation develops a new quantitative-distributive approach to the study of the phonemic structure of word-forms in the Gothic language and their combinatorial models, from the standpoint of which the theoretical provisions of the system-structural organization of language units at all levels and, above all, phonological are supplemented

Through the prism of structural phonology, the work elaborated the ideas of the kinemic theory, according to which the kineme is defined as the smallest articulatory unit, on the one hand, and the smallest perceptual unit, on the other, and the phoneme is considered as a standard block of kinemes for a particular language.

The combinatorics of kinemes is of two types: intraphonemic, realized within the structure of the phoneme simultaneously, and interphonemic, actualized in the combinatorics of adjacent phonemes sequentially over time. The rules of the former and the latter are specific to each particular language and provide clues to revealing the mechanisms of 1) combinability, 2) limited combinability and 3) non-combinability of phonemes.

The Gothic phonemic system reflects the laws of both the Common Indo-European language and the factors that influenced its divergence and were laid in the foundation for the Common Germanic language. This lengthy process took place alongside a change in the grammatical type of language from root isolation through the stage of agglutination to inflection in the pre-writing era. The relics of the features inherent in the first two of these grammatical types have been confirmed in the Gothic manuscripts.

The analysis of the Gothic textual material testifies to the disintegration of the inflectional system and the beginning of the Gothic language system dynamics to the root (analytical) type. Unfortunately, the evolutionary line of the Gothic language was discontinued in the outset of the VIII century AD, and Gothic written records of the IV-VI centuries, which are the oldest among other Old Germanic manuscripts, are available to researchers only in one synchronous state of the language.

The quantitative-distributive approach postulated in the study made it possible to characterize the phonemic structure of the Gothic word-forms and their phonemic combinatorics on the basis of such criteria, as length in phonemes, combinatorics of two classes of phonemes – vowels and consonants – in the initial, medial and final positions.

Based on the results of these characteristics, the research corpus consisting of 9443 Gothic word-forms has been compiled. The textual frequency of these word-forms in the Gothic Bible and minor Gothic manuscripts is 67,464 word-usages.

Eighteen groups of word-forms, according to their length in phonemes, have been identified: the Gothic language macrosystem imposes restrictions on the realization of word-forms longer than nineteen phonemes. Each of the eighteen groups is represented by a different number of word-forms and different textual occurrences of these word-forms in Gothic manuscripts. No eighteen-phoneme word-form has been found.

The analysis has shown that more than 97% of the entire Gothic word-form corpus (9165) is between three and twelve phonemes in length. Five groups of word-forms with a length of five, six, seven, eight and nine phonemes have inessential discrepancies in their frequencies and are the most numerous: of all the registered word-forms, their number is 6849, which is 72.5% of the total (9443) word-form corpus. The number of registered word-forms with the length between one and two and between thirteen and nineteen phonemes is dozens of times smaller and has 278 word-forms, which is 2.9% of the total number of word-forms.

The study of the word-form length in phonemes has shown that the Gothic language macrosystem imposes restrictions on the implementation of one- and two-

phoneme word-forms. Even more constraints of the language system have been imposed on the instantiation of word-forms between 10- and 19-phoneme long: with the increase of length in phonemes, the number of word-forms in a certain class decreases.

An inventory of 656 combinatorial models of the Gothic word-forms, which are canonical forms (CFs), and represent the phonemic structure of the word-form as a chain of realizations of vowel (V) and consonant (C) phonemes, has been established. Theoretically possible number of such models in the Gothic word-forms ranging in length between one and nineteen phonemes is more than one million CFs.

A comparative analysis of the theoretically possible and instantiated canonical forms has shown that the Gothic language exploits its potential by less than one percent. The number of implemented canonical forms utilized by the language decreases with their increasing length. Thus, in the Gothic word-forms with the length ranging between one and six phonemes, the language system employs 50% of its theoretically possible construction potential in one- and two-phoneme word-forms, 62.5% – in three-phoneme word-forms, 68.8% – in four-phoneme word-forms, 59% – in five-phoneme and 51.5% – in six-phoneme word-forms. The realization percentage of theoretically possible word-forms with the increase of their length from seven to nineteen phonemes decreases: for seven-phoneme word-forms it is 32%, for the word-forms consisting of eight-phonemes – 22%, nine-phonemes – 15%, ten-phonemes – 8%, eleven-phonemes – 4,7%, twelve-phonemes – 2.2%. Starting with thirteen-phoneme word-forms, the implementation percentage of theoretically possible phonemic combinatorial types drops sharply and is extremely small: 0.8% in thirteen-phoneme, 0.27% in fourteen-phoneme, 0.05% in fifteen-phoneme, 0.009% in sixteen-phoneme, 0.002% in seventeen-phoneme word-forms. No eighteen-phoneme model of the phonemic structure of word-form in the Gothic language has been registered. The realization percentage of the nineteen-phoneme word-form is 0.0002%.

The inquiry into the realization of combinatorial models of the phonemic structure of word-forms (CFs) in Gothic has shown that it depends on the word-form length in phonemes, on its morphemic structure, on the realization of initial, medial

and final clusters of vowel and consonant phonemes within the word-form, as well as on the configuration of the intraphonemic (nonlinear) combinatorics of the ultimate phonological units – kinemes, which form the inner structure of phonemes as standard, unique and specific to a particular language blocks of kinemes.

The analysis of canonical forms has revealed that the Gothic language does not allow vocalic and consonantal sequences longer than three and four phonemes respectively. Combinability of phonemes in the word-forms of Gothic at the level of the ultimate phonological units is regulated by and depends on both the intraphonemic configuration of kinemes in the internal structure of each phoneme and on the interphonemic combinatorics of, first of all modal, then locational kinemes. The system of phonemes and the system of phonemic oppositions of the Gothic language limit overloading of vocalic and consonantal sequences with marked kinemes.

The role of each canonical form in the language system is determined by its modelling power, which is represented by the frequency of word-forms built on its pattern. The analysis has revealed 24 canonical forms with the length between one and eleven phonemes that have the greatest modelling power. They form the core system and model 50,660 word-form occurrences, which is more than 75% of the word-form corpus total frequency. The core system includes 10 CFs that model 25,440 word-form usages functioning simultaneously as monosyllabic unbound root morphemes and word-forms. The main system, in addition to its core, includes other 47 canonical forms that model 10,152 word-form usages.

The textual frequency of word-forms built on 71 models that make up the main system is 60,812 word-form occurrences and is more than 90% of the total word-form corpus frequency. The remaining 585 canonical forms belong to the periphery and model 6652 word-form usages, which is 9.86% of the registered word-form frequency.

The investigation of 23 out of 656 canonical forms, the phonemic structure of which is not complicated by the implementation of vocalic and consonantal clusters, has shown that they have the largest amount of cumulative modelling power and model 2329 out of 9443 word-forms of the corpus, their textual frequency is 36287 word-form occurrences (53,79% of the total word-form corpus frequency).

The examination of initial vocalic clusters has displayed strict phonological limitations that sharply reduce the phonemic combinability of vowel phonemes in the initial position, especially in the words of Germanic origin. The realization of vocalic clusters in Biblical proper names is the result of transcription of these names from the ancient Greek original.

The actualization of the final vocalic clusters is conditioned mainly by the morphological system of the Gothic language: nominal, pronominal, and verb paradigms. Phonemic combinatorics in the medial vocalic clusters is caused primarily by morphological and morphosyntactic factors.

The treatment of the combinatorial models has identified 33 two-phoneme initial, 75 final and 48 three-phoneme final consonantal clusters registered in the word-forms with textual frequencies of 5038, 5571 and 1315 occurrences, respectively, which are patterned by 13, 141 and 138 canonical forms. Three phonemes: /h^w, z, j/ do not participate at all in the combinatorics of two-phoneme initial and four phonemes/ p, h^w, w, j/ are excluded by the Gothic language macrosystem from patterning in three-phoneme consonantal clusters. Only eight phonemes take part in the formation of four-phoneme final consonantal sequences. The combinability of consonant phonemes in the initial and final consonantal clusters is subject to the principle of sonority hierarchy with its increasing in the initial and decreasing in the final word-form position.

It has been proven for the first time that the Gothic language system prefers covered closed (CVC) and covered open syllables (CV), which is evidenced by their high frequency. Uncovered closed syllables (VC) are typical of the functional parts of speech (particles, prepositions and conjunctions) and for pronouns that were inherited from the Proto-Indo-European and Common Germanic languages.

Canonical types of the Indo-European root phonemic structures were represented by one-syllable morphemes CVC and VC. The root morpheme CVC together with the phoneme of its suffix became a two-syllable stem with the phonemic structure of CV-CV. Uncovered open syllables are realized in the initial position of word-forms.

The frequency of such word-forms is low, since this type of syllable was not typical of the Proto-Indo-European language, it was not characteristic of the Common Germanic language, and as a consequence of this did not receive functional support in the Gothic language.

The sequential realization of vowel and consonant phonemes in uncovered closed syllables and consonant phonemes in closed open syllables was not accompanied by articulatory (pronunciation, anthropophonic) or auditory (perceptual) difficulties, since phonemic combinatorics in such word-forms had no limitations.

The application of quantitative-distributional method with the elements of statistical analysis confirmed the reliability of the obtained results; the accuracy of statistical characteristics is 99.5%.

The proposed methodology can be used in theoretical phonological developments of the Germanic languages and other genetically close and distant Indo-European languages, as well as languages belonging to different grammatical types.

Keywords: distribution, canonical form, kineme, combinatorics, model, word-form, phoneme, phonemic structure, frequency.