

**LEGE ARTIS**

Language yesterday, today, tomorrow

Vol. V. No 1 2020

**PHONEMIC PATTERNING OF WORD-FORMS IN GOTHIC***Seeun Park**Kyiv National Linguistic University, Kyiv, Ukraine*

**Bibliographic description:** Park, S. (2020). Phonemic patterning of word-forms in Gothic. In *Lege artis. Language yesterday, today, tomorrow. The journal of University of SS Cyril and Methodius in Trnava*. Trnava: University of SS Cyril and Methodius in Trnava, 2020, V (1), June 2020, p. 228-274. ISSN 2453-8035

**Abstract:** This study presents a phonological analysis of the word-forms in Gothic at the segmental level. The paper is the first one to establish phonological patterns of the word-forms attested in Gothic manuscripts. Distributional analysis has revealed regularities in the combinability of consonants and vowels within word-forms as well as restrictions and exclusions on them. Statistical analysis proved the systemic nature of phonological patterning in Gothic word-forms.

**Key words:** word-form, phoneme, kineme, phonemic structure, distribution, canonical form.

**1. Introduction**

Studies of the language vocabulary testify to its systematic nature on grammatical and semantic levels (Уфимцева 1962). The statement that vocabulary is not just a set of elements, but an ordered set of interdependent and interrelated elements, that is, a system, must be confirmed not only by grammatical and semantic criteria, but also by phonological ones (Парк 2018a: 4). Linguists have extensively explored the phonemic systems of languages belonging to different grammatical types (Перебийніс 1970; Akidah 2013; Antonsen 1972; Barrack 1998; Fulk 2018; Moradi & Chen 2018; Moulton 1972; Nielsen 2013; Rauch 2017; Trubetzkoy 1969; Voyles 1992) and the phonemic structure of unilateral and bilateral linguistic units, i.e. syllables, morphemes, and words (Васько 2007; Парк 2018b; 2019; Gamkrelidze & Ivanov 1995; Greenberg 1990; Pike 1947; Vennemann 1988).

The linguists who conducted phonological research into different languages, including Germanic, have emphasized the importance not only of establishing phonemic inventories and phonemic opposition systems, but also the need to study the phonemic structure of units at all hierarchical levels of language. The results of the present study testify to the fact that the language system favours one structural model and imposes restrictions on phoneme compatibility in other structural patterns, or completely prohibits instantiation of certain sound shapes consisting of various classes and groups of phonemes.

The discussion of the literature on Gothic phonology concentrates on the most significant works in the field, and does not claim to consider all the findings made before. The developments in the phonology of Gothic do not permit to overview them in the chronological order since there are traces of certain influences of and connections between different linguistic schools. Furthermore, there have always been phonologists, who did not belong to any particular school, but who have been influenced by various theories (Fischer-Jørgensen 1995: xx). The history of linguistic research since its beginning in the 19<sup>th</sup> century has been characterized as "proceeding in two channels, running parallel but differently oriented" (Vachek 1966: 15). In this respect, it is worthwhile to differentiate between phonetics and phonology. The former refers to articulatory, acoustic, and auditory aspects of speech sounds as individual products of individual speakers, actualized every time they are pronounced. Phonology deals with phonemes as language units capable of differentiating meaningful units (morphemes, words, and word-forms) with regard to their function in the language system. According to Krámský, the first scholar to use the word "phoneme" was Dufriche-Desgenettes in 1873 (1974: 21). However, it was Baudouin de Courtenay (1972a: 152), who was the first to give the phoneme its definition as a language unit being "endowed by semasiologized and morphologized functions" in 1894. Most importantly, the concept of the phoneme has become pivotal in the emerging European and American linguistic schools since the thirties of the 20<sup>th</sup> century. These arguments condition the structure of the literature review. The relevant papers based on the

phonetic principle, i.e. making no distinction between the concepts of "sound" and "phoneme", are considered first; a review of phonological works is provided further.

Many scholarly works have been written on the phonology of Gothic. The authors belonging to different linguistic schools bring to light new aspects and offer new insights into the subject. Consider Braune (1883: 1-34), who under the chapter title "Phonology" draws conclusions as to the nature of the Gothic vowel and consonant sounds taking into account the following factors: (1) the origin of the Gothic alphabet; (2) the use of Greek words and Biblical proper names; (3) the transcription of the Gothic proper names in Latin documents and by Latin authors of the 4<sup>th</sup>–8<sup>th</sup> centuries; (4) the testimony of the cognate Germanic languages. Gothic monophthongs and diphthongs in accented and unaccented syllables in the word as well as the subsequent development of the primitive and general Germanic equivalents inherited from the Indo-Germanic vocalic system are described by Wright (1910: 4-83). Gothic consonants are considered as the development of the primitive and general Germanic consonantal system taking into account their correspondences due to the operation of Grimm's Law, Verner's Law, and other consonant changes (*ibid.*).

The first scientific study of the Gothic language has been carried out by Jellinek (1926). Applying comparative analysis to the spelling of ancient Greek words and proper names in Gothic texts, as well as the transcription of Gothic words and proper names in Latin and Greek texts, the spelling of words borrowed into the Romance languages and by establishing sound correspondences of Gothic with other Germanic languages, he describes vowel and consonant sounds (not phonemes), their history, syllabification, accentuation, and combinations of consonant sounds within words. A fair and extensive treatment of Gothic phonetics is suggested by Zadorozhnyj (Задорожний 1960: 49-146), who provides a detailed description of Gothic sounds and different consonantal changes as compared to their Indo-European and Common Germanic correlates. Traditional description of Gothic phonetics is proposed by Agud Aparacio

and Fernandes Alvares (1988: 28-35), who differentiate between long and short vowels, semivowels, occlusive, and fricative sounds.

A more detailed account of different phonological phenomena is given by Gukhman (Гухман 1958: 30-68), who focuses in her research on the Gothic system of writing, reading rules, accentuation, ablaut alternations, spontaneous and combinatorial changes of consonants, which took place before the time of the Gothic written records. The system of Gothic phonemes is described in terms of classical phonology by Mossé (1969: 51-69) with the emphasis on the classification and pronunciation of short and long vowels, diphthongs, sound values of diagraphs, consonants, accentuation, grammatical alternations according to Verner's Law, assimilation, dissimilation, simplification of consonantal clusters; much attention is paid to the phonetic changes in the unaccented syllables.

Another landmark in the Gothic phonological studies is the work of Marchand (1973), in which the author, aiming "to determine the pronunciation of Gothic", uses such criteria, as "(1) the origin of the alphabet; (2) loan words and transcriptions of proper names; (3) internal evidence; (4) comparative evidence" (op. cit., 12). Marchand describes the pronunciation of the 4<sup>th</sup> century Greek, certain Latin, and runic signs (op. cit., 27-30); provides orthographical deviants explaining them as different types of errors: mechanical errors, errors, which may indicate pronunciation (op. cit., 37-57); discusses phonemic oppositions (op. cit., 60-64); considers comparative evidence afforded by other Germanic and Indo-European languages (op. cit., 79-101). One of the most important contributions to Gothic phonology has been made by Bennett (1980), who outlines phonological history of Gothic, explains essentials of phonologic and analogic changes, developments of short vowels in originally medial and final syllables.

Statistical data of Gothic initial and final consonants and their clusters as the ratio of two frequencies – the average textual frequency and lexicon (dictionary) frequency –

are provided by Joos (1942). Snædal (2009) presents various statistics concerning East Germanic, i.e. Gothic morphology and individual characters (letters and numerals), and principal speech sounds (phonemes) in the Gothic manuscripts (2013). Rauch (2011) in terms of generative phonology establishes Gothic consonantal (op. cit., 47-49) and vocalic (op. cit., 59-60) inventories; she describes phonological constraints in consonants (op. cit., 41-46) and vowels (op. cit., 51-58); and under the subtitle "Phonological architecture" segments the Gothic word "into a hierarchy of descending tiers: the foot, the syllable, the skeletal or CV tier and the root tier" (op. cit., 35-40); gives matrices of phonetic distinctive features for the Gothic consonants and vowels (ibid.).

The phonological evolution of the Old Germanic languages, including Gothic, is presented within the framework of "kinemic theory" by Plotkin (2008: 33-34). This theory views the system of phonemes and phonemic oppositions in a dynamic aspect, i.e. as the process, in which particular sound changes are described as links in one chain of causes and effects that underlie the phonological history of every single language under investigation. In his earlier writings, Plotkin (1978) uses the term "kinakeme" coined by Baudouin de Courtenay (1972c: 280) to designate the ultimate phonological unit of language. This term is a blend composed of *κίνημα* "kineme" – the elementary unit of articulatory work, and *ακουσμός* "acousmeme" – the elementary unit of auditory perception, respectively (1972b: 325). The author of the kinemic theory (Plotkin 2008: 26) explains:

*"Of these three terms, 'kineme' appears to be the most suitable term for the ultimate phonological unit. <...> As for the term 'kinakeme', which was previously used by the present author to designate the ultimate language unit, it appears to be unnecessarily complicated by the insertion of a reference to auditory perception. <...> the reference to the initial action is quite sufficient".*

Linear syntagmatics of the ultimate phonological units in the initial, medial, and final consonantal groups of Gothic was investigated by Vasko (Васько 1997). In his research, the scholar establishes the frequency, regularities, and restrictions in the consonantal clusters of Gothic lexemes selected from Feist (1939), analyzes

interphonemic combinability of the ultimate phonological units, which are very productive in Gothic and based mainly on phonological contrasts.

Yet, in spite of the extensive scholarship done since the inception of Gothic studies, some issues of Gothic phonology still remain the subject of discussions in Germanic linguistics. An inquiry into the Gothic phonemic system constructional potential and its implementation, preferences, limitations, and prohibitions on phonemic patterning, as well as statistical investigation of the corpus under analysis merit further consideration.

This article aims to establish the inventory of the word-forms construction patterns, explore their potential, frequency, and functional load; identify and examine systemic and peripheral models, investigate the vowel and consonant phonemes patterning, reveal regularities, constraints, and exclusions on the phonemic combinability, and define the average word-forms length and its mean square deviation, estimate the research exactitude, and other data. Neither of these topics has been studied before. The aim of the study is achieved by applying the methods of quantitative, distributional, and statistical analyses to the phonological structure of word-forms attested in the Gothic Bible and minor Gothic manuscripts. Obtained by mechanical sampling, 9443 word-forms presented by Tollenaere and Jones (1976), with the exception of word fragments, number symbols, and emended word-forms, have made the corpus of the present study. The word-forms sampling frequency is 67464. Each of the registered word-forms has also been cross-checked in Köbler (1989), Snædal (1998), and Streitberg (2000).

## **2. The phonemic system of Gothic**

Before embarking on the analysis of phonemic patterning in the Gothic word-forms, it is necessary to decide what phonemic system of the Gothic language will be taken for the research. In this study, the subsystem of Gothic consonant phonemes proposed by Vasko (Васько 1997) and the subsystem of Gothic vowel phonemes presented by

Plotkin (2008) are used. These subsystems are logically and linguistically substantiated and do not contradict each other as far as the phonological theory is concerned.

Furthermore, it is necessary to specify the approach to the notion of phoneme, which, from its inception, was introduced to denote the ultimate indivisible unit of the language system. However, the divisibility of phonemes into smaller phonologically relevant entities charged with semiotic value became evident to the representatives of the linguistic school of Prague as early as in the mid-thirties of the 20<sup>th</sup> century (Trubetzkoy 1969: 67; Vachek 1976b: 17); later research brought to light "distinctive features" as a universal inventory assumed to be used by all human languages for building phonemes as bundles of such features (Jakobson 1971: 422, 425, 484-486). With the development of phonological theory, other terms, such as "mérisme" (Benveniste 1966: 121) meaning "delimitation" in Greek and "phononeme" (Grucza 1970: 77) were coined and introduced, but neither of them seems acceptable "since the ultimate phonological unit is not in itself a unit of sound" (Plotkin 2008: 25). The concept of "distinctive feature" was criticized for its imprecision (Fischer-Jørgensen 1995: 146-147) since it did not reflect its "emic" status as a linguistic unit. The term "kineme" proposed by Plotkin (2008: 26) appears to be the most appropriate as it properly reflects the essence of the ultimate "emic" unit of sound generating movement, and it will be used in this sense here below. Plotkin (op. cit., 63) explains:

*"A block of kinemes constituting a phoneme is a standard unity providing for automatic materialization of its underlying substance – a systemically organized block of several aggregates for phonatory actions. Materialization of a single kineme is automatic in the sense that the neural impulse it represents in the language system activates the aggregate of actions as an integral unity, a complex formed in the process of mastering the sound pattern of the language".*

Thus, the phoneme is considered as a language-specific standard block of kinemes, while Jakobsonian (Jakobson 1971: 420-421) and Chomskyan (Chomsky & Halle 1991: 299-300) "distinctive features" are universal, mere inventories, and represent "the phonetic capabilities of man" (ibid.). Kinemes, on the contrary, have all the attributes

of linguistic units: they are paradigmatically organized, syntagmatically combined, and they carry semantic load (Plotkin 2008: 45, 52, 61).

Consonantal modal kinemes (Plotkin 2008: 82-86) determine the nature of the obstacle to the air stream and are represented in Gothic by the sonority opposition for controlling the work of the vocal cords and the striction opposition that deals with other obstacles. Consonantal locational kinemes (op. cit., 87-88) determine the choice of the active and passive articulatory organs, thus forming in Gothic the opposition of labiality (pre-centrality according to the active organ of speech), the opposition of palato-alveolarity / velarity / pharyngeality (post-centrality according to the active organ of speech, and the opposition of post-dentality (post-centrality according to the passive organ of speech). Vocalic modal kinemes (op. cit., 71-73) control tongue elevation (low vs. non-low or high vs. non-high), while vocalic locational kinemes (op. cit., 73-75) control timber characteristics dependent on the horizontal tongue position and lip movement (front vs. non-front or labialized vs non-labialized).

Within the framework of the kinemic theory, any phoneme can be described as a bundle of kinemes, e.g., the Gothic phoneme /l/ is defined as a standard block of kinemes consisting of the positive modal kineme of striction, negative modal kineme of obstruence, positive modal kineme of sonority, and locational negative kinemes of labiality, palato-alveolarity, and post-dentality.

### **3. The phonemic structure of the Gothic word-forms at the segmental level**

The study of phonemic patterning in the Gothic word-forms involves determining the length of each registered word-form in phonemes, establishing the inventory of models (patterns) of the phonemic structure of word-forms, calculating the frequency of each model, applying statistical methods to verify the reliability of the study, identifying clusters of vowel and consonant phonemes in the structure of word-forms, analyzing the mechanism of their patterning in order to expose factors, preferable patterns,



restrictions, and prohibitions on combinability within the phonemic structure of word-forms.

### *3.1 The length of the Gothic word-forms in phonemes*

One of the important structural features of a word-form is its length in phonemes. Torsuyev notes that "the possible number of phonemes in a word / word-form itself is interesting in the structural characterization of a word / word-form" (Торсуев 1962: 5). Smirnitskij defined the word-form in the following way: "The word-form is a certain word in a certain grammatical form and of a certain phonemic shape, by which one particular lexical meaning is expressed" (Смирницкий 1954: 11). This section of the study focuses on bringing to light the registered Gothic word-form length in phonemes. The corpus of the word-forms under analysis is 9443; the textual frequency of the word-forms is 67464.

As a result of the corpus phonemic structure analysis, word-forms from one to nineteen phonemes in length have been registered. No eighteen-phoneme word-forms have been found. Thus, the whole corpus of word-forms is divided into eighteen groups. Each of the eighteen groups is represented by a different number of word-forms and their different textual frequency in the Gothic manuscripts under study (See Table 1).

Table 1. Distribution of the phonemic word-form length in Gothic

#	Number of phonemes in the word-form	Number of the registered word-forms	% of the word-form usage in the sample	Absolute frequency of the word-form	% of the word-form usage from their absolute frequency
1.	1	2	0.02	658	0.97
2.	2	53	0.56	11175	16.56
3.	3	200	2.12	13032	19.31
4.	4	619	6.55	9512	14.10
5.	5	1175	12.44	10191	15.10
6.	6	1532	16.22	6981	10.34
7.	7	1548	16.39	5634	8.35
8.	8	1428	15.12	3914	5.80
9.	9	1166	12.35	3002	4.44
10.	10	724	7.66	1473	2.18
11.	11	521	5.51	1058	1.56
12.	12	252	2.67	477	0.71
13.	13	128	1.35	205	0.30
14.	14	64	0.67	109	0.16
15.	15	20	0.21	29	0.04

16.	16	6	0.06	8	0.01
17.	17	3	0.03	4	0.006
18.	18	<i>Not registered</i>			
19.	19	2	0.02	2	0.003
		<b>Total: 9443</b>	<b>Total: 99.95%</b>	<b>Total: 67464</b>	<b>Total: 99.94%</b>

**One-phoneme word-forms.** This group is represented by 2 word-forms. Their textual frequency is 658, e.g., Goth. *ei* /i/ – partic., conj. "in that case, under those circumstances, then, thereby, such that, a time wherein, such manner as, in a way that" (Köbler 1989: 131-133; Snædal 1998: 205-213; Streitberg 2000: 53; Tollenaere & Jones 1976: 43-44).

**Two-phoneme word-forms.** This group is represented by 53 word-forms. Their textual frequency is 11175, e.g., Goth. *in* – prep. "in, into, unto, onto, at, within, when, while, in virtue of, because of, for the sake of, through, due" (Köbler 1989: 294-300; Snædal 1998: 440-470; Streitberg 2000: 125; Tollenaere & Jones 1976: 91-96).

**Three-phoneme word-forms.** This group is represented by 200 word-forms. Their textual frequency is 13032, e.g., Goth. *þan* – adv., conj. "then, at that time, but, this time, thereupon, thereafter, now then, furthermore, moreover, when, thus, whenever" (Köbler 1989: 535-538; Snædal 1998: 1042-1051; Streitberg 2000: 17; Tollenaere & Jones 1976: 176-178).

**Four-phoneme word-forms.** This group is represented by 619 word-forms. Their textual frequency is 9512, e.g., Goth. *þata* – nom. and acc., neut., sing. of demonstr. pron. "that" (Köbler 1989: 541; Snædal 1998: 860-866; Streitberg 2000: 127; Tollenaere & Jones 1976: 180-183).

**Five-phoneme word-forms.** This group is represented by 1175 word-forms. Their textual frequency is 10191, e.g., Goth. *ahman* – noun, acc., neut., sing. *ahma* "spirit" (Köbler 1989: 15; Snædal 1998: 20; Streitberg 2000: 103; Tollenaere & Jones 1976: 8).

**Six-phoneme word-forms.** This group is represented by 1532 word-forms. Their textual frequency is 6981, e.g., Goth. *kunnub* – 2<sup>nd</sup> pers., sing., pret. ind. of pret.-pres. verb *kunnan* "be acquainted, to know, to understand" (Köbler 1989: 347; Snædal 1998: 618; Streitberg 2000: 47; Tollenaere & Jones 1976: 119).

**Seven-phoneme word-forms.** This group is represented by 1548 word-forms. Their textual frequency is 5634, e.g., Goth. *armands* – Part. I, nom., masc., sing. of *armands* "having mercy on" (Köbler 1989: 58; Snædal 1998: 95; Streitberg 2000: 241; Tollenaere & Jones 1976: 23).

**Eight-phoneme word-forms.** This group is represented by 1428 word-forms. Their textual frequency is 3914, e.g., Goth. *galaubeinai* – noun, dat., fem., sing. of *galaubeins* "belief, faith" (Köbler 1989: 199-200; Snædal 1998: 292-294; Streitberg 2000: 233; Tollenaere & Jones 1976: 61).

**Nine-phoneme word-forms.** This group is represented by 1166 word-forms. Their textual frequency is 3002, e.g., Goth. *insandida* – 1<sup>st</sup> and 3<sup>rd</sup> pers. sing., pret. ind. of the weak verb I *insandjan* "send, dispatch" (Köbler 1989: 304; Snædal 1998: 913-914; Streitberg 2000: 77; Tollenaere & Jones 1976: 97).

**Ten-phoneme word-forms.** This group is represented by 724 word-forms. Their textual frequency is 1473, e.g., Goth. *fauramableis* – noun, nom. and gen., masc. *faúramableis* "ruler, prince, chief, head man, governor, official" (Köbler 1989: 146; Rousseau 2016: 97; Snædal 1998: 230; Streitberg 2000: 125; Tollenaere & Jones 1976: 47).

**Eleven-phoneme word-forms.** This group is represented by 521 word-forms. Their textual frequency is 1058, e.g., Goth. *galaubidedun* – 3<sup>rd</sup> pers. pl., pret. ind. of the weak verb I *galaubjan* "to believe, to permit, to have faith in, to be confident of, to have confidence in" (Köbler 1989: 199-200; Snædal 1998: 636-638; Streitberg 2000: 49;

Tollenaere & Jones 1976: 62).

**Twelve-phoneme word-forms.** This group is represented by 252 word-forms. Their textual frequency is 477, e.g., Goth. *swnagogafada* – noun, dat., masc., sing. of *swnagogafaps* "synagogue leader, ruler of the synagogue" (Köbler 1989: 518; Snædal 1998: 1009; Streitberg 2000: 183; Tollenaere & Jones 1976: 170).

**Thirteen-phoneme word-forms.** This group is represented by 128 word-forms. Their textual frequency is 205, e.g., Goth. *gaswiltandans* – Part. I, nom., masc., pl. of *gaswiltandans* "those who cease living" (Köbler 1989: 224-225; Snædal 1998: 1007-1008; Streitberg 2000: 275; Tollenaere & Jones 1976: 68).

**Fourteen-phoneme word-forms.** This group is represented by 64 word-forms. Their textual frequency is 109, e.g., Goth. *sildaleikidedun* – 3<sup>rd</sup> pers. pl., pret. ind. of the weak verb I *sildaleikjan* "to marvel, to wonder, to be amazed" (Köbler 1989: 475-476; Snædal 1998: 913-914; Streitberg 2000: 15; Tollenaere & Jones 1976: 159).

**Fifteen-phoneme word-forms.** This group is represented by 20 word-forms. Their textual frequency is 29, e.g., Goth. *ufarfulljandans* – Part. I, nom., masc., pl. of *ufarfulljandans* "those who are overfilling, those who are filling to superabundance" (Köbler 1989: 562; Snædal 1998: 268; Streitberg 2000: 277; Tollenaere & Jones 1976: 191).

**Sixteen-phoneme word-forms.** This group is represented by 6 word-forms. Their textual frequency is 8, e.g., Goth. *mipanakumbidedun* – 3<sup>rd</sup> pers. pl., pret. ind. of the weak verb I *mipanakumbjan* "recline together with to eat, lie down to a meal together" (Köbler 1989: 392; Snædal 1998: 616; Streitberg 2000: 15; Tollenaere & Jones 1976: 134).

**Seventeen-phoneme word-forms.** This group is represented by 3 word-forms. Their

textual frequency is 4, e.g., Goth. *ufarhiminakundans* – adj., nom., masc., pl. of *ufarhiminakunds* "born of heaven above, heavenly, of heavenly origin" (Köbler 1989: 563; Snædal 1998: 1100; Streitberg 2000: 275; Tollenaere & Jones 1976: 191).

**Nineteen-phoneme word-forms.** This group is represented by 2 word-forms. Their textual frequency is 2, e.g., Goth. *gahpanmipsandidedum* – 1<sup>st</sup> pers. pl., pret. ind. of the weak verb I *gamipsandjan* "and then we sent with" – (Köbler 1989: 209; Miller 2019: 198, 268, 670; Rauch 2011: xxi; Snædal 1998: 914-915; Streitberg 2000: 311; Tollenaere & Jones 1976: 60, 70).

As shown in Table 1, 9165 Gothic word-forms with a length from three to twelve phonemes comprise more than 97% of the sample, whereas the number of the registered word-forms with a length of one to two and of 13 to 19 phonemes is more than 60 times smaller: there are 278 word-forms, which is 2.9% of the total number of word-forms. According to the laws of the Gothic language, only two vowel phonemes /o/ and /i/, graphically represented by *o* and *ei*, respectively, form one-phoneme word-forms. The initial position in the Gothic word-forms is occupied by all 19 consonant and 7 vowel phonemes with the only restriction for the phoneme /j/ to occupy the word-form final position.

### 3.2 Statistical analysis

Data on the number of phonemes that can occupy word-form initial and final positions, i.e. 26 and 25 phonemes, respectively, allow calculating the theoretically possible number of two-phoneme word-forms:  $26 \times 25 = 650$ . In the corpus under analysis, only 53 two-phoneme word-forms are registered. A comparison of the theoretically possible number with the registered quantity of instantiated models of two-phoneme word-forms (650 and 53, respectively) leads to the conclusion that the Gothic language macrosystem imposes restrictions on the implementation of one- and two-phoneme word-forms. Calculations presented in Table 1 concern the implementation of word-forms with the length of 10 to 19 phonemes, show even more limitations of the Gothic

language macrosystem: as the word-form length increases, the number of word-forms decreases sharply.

To calculate the average phonemic length of the registered word-forms, the formula proposed by Perebyjnis (Перебийніс 2002: 38) is used:

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i}, \text{ in which:}$$

$\bar{x}$  – average word-form length in phonemes;

$\sum$  – total number of the registered word-forms;

$x_i$  – number of phonemes in the word-form;

$n_i$  – number of word-forms of a certain length in phonemes;

$\sum n_i$  – total number of word-forms of all lengths in phonemes.

After performing the necessary operations, the average phonemic length is calculated:

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i} = \frac{70045}{9443} = 7.42 \text{ phonemes.}$$

Thus, the average word-form length in the corpus of the Gothic word-forms under analysis is 7.42 phonemes. This value depends on the number of word-forms of each length: an increase or decrease in the number would change the average length of the word-form. When calculating the average length of the word-form, it is essential to set the limits of oscillation of both average and absolute lengths by the formula of the mean

square deviation:  $\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 n_i}{\sum n_i}}$  (ibid., 48).

Having made all the calculations, the mean square deviation is found by the above-given formula, in which:

$\sigma$  – mean square deviation;

$\bar{x}$  – average word-form length in phonemes;

$x_i$  – number of phonemes in the word-form;

$n_i$  – number of word-forms of a certain length in phonemes;

$\sum$  – total number of the registered word-forms;

$\sum n_i$  – total number of word-forms of all lengths in phonemes.

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2 n_i}{\sum n_i}} = \sqrt{\frac{52050.4652}{9443}} = \sqrt{5.512068} = 2.3477 = 2.35.$$

To verify the accuracy of the obtained results, it is necessary to estimate the oscillation measure of the average word-form length  $\sigma_{\bar{x}}$ , which is calculated by the following formula  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\sum n_i}}$  (ibid., 51).

Having used all previously obtained data, the measure of oscillation is set:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{\sum n_i}} = \frac{2.3477}{\sqrt{9443}} = \frac{2.3477}{97.1750} = 0.0241 = 0.02.$$

The mean square deviation  $\sigma = 2.35$ , and  $\sigma_{\bar{x}} = 0.02$ . In the range from  $\bar{x} - 2\sigma_{\bar{x}}$  to  $\bar{x} + 2\sigma_{\bar{x}}$ , i.e.  $7.42 - 0.02$  to  $7.42 + 0.02$  ( $7.40 - 7.44$ ), there are possible oscillations in the average Gothic word-form length that are caused by statistical factors.

Since word-form fragments, word-forms with spaces, letters denoting numbers, word-forms with individual letters restored by Streitberg (2000) were not registered in the sample, the value of  $\bar{x}$  cannot be calculated precisely. According to Perebyjnis (Перебийніс 2002: 60-61):

*"...  $\bar{x}$  and other statistical characteristics... are calculated with a certain amount of error, that is, with a certain deviation of the value that we would obtain if we survey the whole corpus. This inaccuracy is called a margin of error (or inaccuracy) of the study" and is calculated by the formula  $\varepsilon = \frac{K\sigma_{\bar{x}}}{\bar{x}}$ , in which:  $K$  is a constant that determines the confidence coefficient and equals to 1.96'."*

Relative margin of error in the study is found according to the formula:

$\varepsilon = \frac{K\sigma_{\bar{x}}}{\bar{x}} = \frac{1.96 \times 0.02}{7.42} = \frac{0.0392}{7.42} = 0.0052 = 0.005$ . This means that the value of  $\bar{x}$  (average phonemic length of the registered word-forms in Gothic) in this study is calculated with a relative error of 0.5%, and the exactitude of the statistical characteristics is 99.5%. Consequently, statistical data of the phonemic patterning of the word-forms in Gothic are highly precise.

#### 4. Distributional analysis at the phonemic level

The phonemic combinatorial model in a word-form is represented by a canonical form (CF), where **C** is any consonant phoneme and **V** is any vowel phoneme. The consonantal Gothic subsystem consists of 19 phonemes, the vocalic system – of seven phonemes. The theoretical number of word-forms of a certain length in phonemes is calculated by the formula  $\Sigma = 2^n$ , where  $n$  is the number of phonemes that structure the word-forms (length of word-forms in phonemes) (Перебийніс 1970: 160). Thus, for nineteen-phoneme word-forms, the number of canonical forms is  $\Sigma = 2^{19}$ , that is 525288 CFs, etc.

##### 4.1 Canonical forms

It is quite certain that not all theoretically possible canonical forms (combinatorial models) of the phonemic structure of the Gothic word-forms can be instantiated: like in any other language, it depends on many factors: 1) the number of phonemes in a word-form; 2) the morphemic structure of a word-form; 3) the implementation of consonantal and vocalic clusters of phonemes within a word-form in different (initial, medial, and final) positions; 4) the phoneme that occupies initial position in a word-form, and 5) the kinemic structure of the phonemes that constitute a word-form.

Table 2 shows the number of theoretically possible and registered canonical forms; instances of the Gothic word-forms of each particular length are provided in the right-side column. The table shows that percentage of the canonical forms instantiated in word-forms of each particular phonemic length depends on the number of phonemes constituting the word-forms, ranging from 68.75% (the highest measure of realization) in 4-phoneme word-forms to 0.0002% (the lowest measure of realization) in 19-phoneme word-forms. After analyzing the corpus of 9443 Gothic word-forms, their inventory of 656 canonical forms (CFs) was established.

As shown in Table 2, the Gothic language does not use all theoretically possible canonical forms from 1 to 19 phonemes long. The longer the word-forms in phonemes



are, the smaller the percentage of realization of theoretically possible canonical forms by the language is, and the smaller the number of word-forms modelled by these canonical forms is.

Table 2. Number of theoretically possible and registered canonical forms compared

Number of phonemes in a word-form	Theoretically possible number of canonical forms	Number of the registered canonical forms	% of realization	Canonical form (example from the corpus under analysis)
1	2	1	50	V – o – interj. "o, oh, aha, ha";
2	4	2	50	VC – ip – conj "but, however, in any case";
3	8	5	62.5	CVC – her – adv. "here, hither";
4	16	11	68.75	CVCC – barn noun, nom., neut. "child";
5	32	19	59.38	CCVCV – staina – noun, dat. masc. "stone, rock";
6	64	33	51.56	CVCCVC – sunjus– noun, gen., fem. "truth";
7	128	41	32.03	CVCCCCV – waurstwa – noun, nom., masc. "worker";
8	256	57	22.27	CCVCCVCV – swiknipai – noun, dat. fem. "purity, chasity";
9	512	77	15.04	CVCCVCCVC – gaggandin – Part. I, dat., masc., sing. of irreg. verb gaggan "to go, to walk";
10	1024	81	7.91	CVCVCVCCCV – fauragaggja – noun, nom., masc. sing. "superintendent";
11	2048	96	4.69	CVCVCVCCCCV – manamaurprja – noun, nom., masc., sing. "murderer";
12	4096	91	2.22	CVCCVCCVCCVC – midjungardis – noun, gen., masc. sing. "the habitable world, earth";
13	8192	71	0.87	CVCVCVCVCCVCV – himinakundana – adj., acc., masc. sing. "heavenly, celestial";
14	16384	44	0.27	CVCVCVCVCVCCVC – witodalaisarjos – noun, nom., masc., pl., "law teacher, legal exegete, scribe";
15	32768	17	0.052	CVCCVCVCCVCCVCV – sildalikjandona – Part. I, nom., pl., neut., of the weak verb I sildaleikjan "marvel, wonder";
16	65536	6	0.0092	CVCCVCVCCCVCCVCV – gaswikunþjandona – Part. I, pl. neut., of the weak verb I gaswikunþjan "make manifest, make openly known";
17	131072	3	0.0023	CVCVCVCVCCCVCCVCV – miþanakumbjandane – Part. I, gen., pl. masc. of the weak verb I miþanakumbjan "lie down to a meal together";
18	262144	<b>Not registered</b>		
19	524288	1	0.0002	CVCCVCCVCCVCCVCVCVC – gahþanmiþsandidedum – 1 pers. pl. pret. of the weak verb I gamiþsandjan "to send / dispatch together with".
<b>Total:</b>	<b>1048574</b>	<b>656</b>	<b>0.0625</b>	

The analysis shows that the use of theoretically possible canonical forms with a length from one to six phonemes is high (from 68.75% to 50%). As the length of the canonical

form increases, the percentage of its realization from the theoretically possible number decreases: after the 6-phoneme canonical forms it is less than 50%, and it decreases as the length of the canonical forms increases (from 32.03% in 7-phoneme to 0.0002% in 19-phoneme canonical forms). The study of the phonemic structure of word-forms shows that the number of canonical forms registered in the Gothic language makes an extremely small portion of their theoretically possible number: 656 out of 1048574, which is 0.0625%.

It is necessary not only to establish the inventory of the canonical forms used in the language, but also to determine the role that each canonical form plays in it. All the Gothic corpus of 67464 word-form usages is patterned according to 656 canonical forms (CFs), each of them having different frequency, and consequently, being characterized by various modelling power (Папк 2018a: 6). Thus, there are CFs, such as **CVC**, e.g., Goth. *jah* – conj. "and, but, also", **VC**, e.g., Goth. *ut* – adv. "out, outside", **CV**, e.g., Goth. *bi* – prep. "by, at, near", etc. with very high frequency of 9588, 5948, and 5227 of usages, which is 14.212%, 8.816% and 7.748%, respectively, of the total word-forms frequency; and there are such CFs, as **CCVCVCCVCV**, e.g., Goth. *swikunþaba* – adv. "plainly, clearly, in an obvious manner", **CVCVCVCVCVCV**, e.g., Goth. *fauragahugida* – 3<sup>rd</sup> pers. sing., pret. ind. of the weak verb I *fauragahugjan* "make up one's mind beforehand", and **VCCCVCCCV**, e.g., Goth. *andbahtjaina* – 3<sup>rd</sup> pers. pl., pres. opt. of the weak verb I *andbahtjan* "serve, administer, perform", with much lower frequency of 25, 15, and 5, respectively. Such canonical forms, as **VCVCCCVCCC**, e.g., Goth. *ufarskafts* – "altar, first fruit", **CCVCVCCCC**, e.g., Goth. *frawaurhts* – "offence, transgression, evil-doing", and **VCVCCVCVCVCVCVCVCC**, e.g., Goth. *ufarhiminakundans* – adj., pl., nom, masc., weak decl. of *ufarhiminakunds* "born of heaven above, of heavenly origin, heavenly", are represented each by a single word-form.

For further study of the phonemic distribution in Gothic word-forms, it is necessary to determine, which canonical forms comprise the core system, the main system or centre,

and periphery in terms of Prague school of linguistic thought (Daneš 1966; Vachek 1966: 27; 1976a), which is 75%, 90%, and 10%, respectively. Table 3 shows 24 canonical forms (CFs) with the highest modeling power in the decreasing frequencies order. These CFs represent the core system, they model 50660 Gothic word-forms, which comprise 75.091% of all the sample.

Table 3. Modelling power of the most frequent canonical forms in Gothic

#	Canonical form	Example of a word-form and its translation	Modelling power (frequency)	Cumulative modelling power (frequency)	Modelling power in %	Cumulative modelling power in %
1.	<b>CVC</b>	<i>gam</i> – "came"	9588	9588	14.212	14.212
2.	<b>VC</b>	<i>ut</i> – "out, outside"	5948	15536	8.816	23.028
3.	<b>CV</b>	<i>bi</i> – "by, at, near"	5227	20763	7.748	30.776
4.	<b>CVCV</b>	<i>runa</i> – "secret plan"	4803	25566	7.119	37.895
5.	<b>CVCVC</b>	<i>sunus</i> – "son"	4087	29653	6.058	43.953
6.	<b>CVCCV</b>	<i>manna</i> – "man, person"	2386	32039	3.536	47.490
7.	<b>VCCVC</b>	<i>unsar</i> – "our"	2096	34135	3.107	50.597
8.	<b>CVCCVC</b>	<i>wiljan</i> – "to wish"	1976	36111	2.929	53.526
9.	<b>VCCV</b>	<i>ahtau</i> – "eight"	1813	37924	2.687	56.213
10.	<b>CVCC</b>	<i>barn</i> – "child"	1661	39585	2.462	58.675
11.	<b>CVCVCV</b>	<i>rodida</i> – "spoke"	1493	41078	2.213	60.888
12.	<b>VCV</b>	<i>augo</i> – "eye"	1349	42427	1.999	62.888
13.	<b>VCC</b>	<i>ains</i> – "one, alone"	1255	43682	1.860	64.748
14.	<b>CVCVCVC</b>	<i>laisareis</i> – "teacher"	1221	44903	1.810	66.558
15.	<b>V</b>	<i>ei</i> – "in that case"	658	45561	0.975	67.533
16.	<b>CCVCVC</b>	<i>broþar</i> – "brother"	652	46213	0.966	68.500
17.	<b>VCCVCVC</b>	<i>aggilus</i> – "angel"	609	46822	0.902	69.403
18.	<b>CCVCCV</b>	<i>stibna</i> – "voice"	607	47429	0.899	70.302
19.	<b>CVCVCC</b>	<i>mikils</i> – "great, large"	581	48010	0.861	71.163
20.	<b>CVCVCCVC</b>	<i>bokarjos</i> – "scribes"	577	48587	0.855	72.019
21.	<b>CVCVCCV</b>	<i>fijapwa</i> – "enmity"	564	49151	0.836	72.855
22.	<b>VCCVCV</b>	<i>idreiga</i> – "repentence"	507	49658	0.751	73.606
23.	<b>CVCVCVCV</b>	<i>samaleiko</i> – "similarly"	506	50164	0.750	74.356
24.	<b>CVV</b>	<i>saei</i> – "the one who"	496	50660	0.735	75.091
<b>TOTAL:</b>			<b>50660</b>	<b>50660</b>	<b>75.091</b>	<b>75.091</b>

As shown in Table 3, there are four three-phoneme CFs with the highest cumulative modelling power: 12688 word-form usages, which is 18.807% of the total frequency, two two-phoneme CFs with cumulative modelling power of 11175 word-form usages, which is 16.564% of the total frequency, three five-phoneme CFs with cumulative modelling power of 8569 word-form usages, which is 12.701% of the total frequency, three four-phoneme CFs with cumulative modelling power of 8277 word-form usages, which is 12.269% of the total frequency, six six-phoneme CFs with cumulative

modelling power of 5816 word-form usages, which is 8.621% of the total frequency, two seven-phoneme CFs with cumulative modelling power of 2394 word-form usages, which is 3.548% of the total frequency, two eight-phoneme CFs with cumulative modelling power of 1083 word-form usages, which is 1.605% of the total frequency, and one one-phoneme CF with the frequency of 658 word-form usages, which is 0.975% of the total frequency. The main system together with the core of the system include 71 canonical forms (10.823% of all the CFs) that model 60812 word-form usages, which is 90.14%. The rest 585 canonical forms model 6652 word-form usages, which is 9.86% of the word-forms corpus under analysis, which represents the Gothic language word-forms periphery.

#### *4.2 Gothic two-phoneme initial vocalic clusters*

In the Gothic word-form corpus under analysis, 61 word-forms containing two-phoneme vocalic clusters in the initial position were attested. This is 0.65% of the total word-form number (9443). Their frequency is 767 word-form usages, which is 1.14% of the total word-form frequency (67464). Among these word-forms, only seven are of Germanic origin with the frequency of 25 word-form usages. They all begin with a two-phoneme vocalic cluster *iu* /iu/, which represents the only Gothic biphonemic diphthong with no signs of being monophonemicized (Ebbinghaus 1971; Plotkin 2008: 127), e.g.:

- 1) Goth. *iumjons* – noun, nom., fem. pl. of *iumjo* "crowd, throng" (Köbler 1989: 317; Snædal 1998: 516; Streitberg 2000: 11; Tollenaere & Jones 1976: 104);
- 2) Goth. *iup* – adv. "up, upwards" (Köbler 1989: 317; Snædal 1998: 516; Streitberg 2000: 59; Tollenaere & Jones 1976: 104);
- 3) Goth. *iupa* – adv. "on high, above, aloft, high up" (Köbler 1989: 317-318; Snædal 1998: 516; Streitberg 2000: 458; Tollenaere & Jones 1976: 104);
- 4) Goth. *iupana* – adv. "from above, from before, again from the start, anew" (Köbler 1989: 318; Snædal 1998: 516; Streitberg 2000: 359; Tollenaere & Jones 1976: 104);
- 5) Goth. *iupapro* – adv. "from above" (Köbler 1989: 318; Snædal 1998: 516; Streitberg 2000: 29; Tollenaere & Jones 1976: 104);

6) Goth. *iusila* – noun, nom., fem. sing. "improvement, amelioration, betterment, ease" (Köbler 1989: 318; Snædal 1998: 517; Streitberg 2000: 405; Tollenaere & Jones 1976: 104);

7) Goth. *iusiza* – adj. (supplet. of *gops* "good"), nom., masc. sing. "better, superior" (Köbler 1989: 318; Snædal 1998: 517; Streitberg 2000: 357; Tollenaere & Jones 1976: 104).

In addition to these, 54 word-forms of the corpus with the frequency of 742 word-form usages represent proper names of the Biblical personages, geographical names, and names of nationalities, which were transcribed from the Old Greek original, the so called "*Greek Vorlage*" of the New Testament (Miller 2019: 18-20), e.g.:

1) *Iesus* /iesus/ – "Jesus" (Snædal 1998: 420; Streitberg 2000: 81; Tollenaere & Jones 1976: 85);

2) *Iakob* /iakob/ – "Jakob" (Snædal 1998: 416; Streitberg 2000: 231; Tollenaere & Jones 1976: 85);

3) *Iairusalem* /ierusalem/ – "Jerusalem" (Snædal 1998: 415; Streitberg 2000: 93; Tollenaere & Jones 1976: 85);

4) *Iaurdanu* /iordanu/ – noun, acc., masc. sing. "the Jordan river" (Snædal 1998: 417; Streitberg 2000: 55; Tollenaere & Jones 1976: 85);

5) *Iudaialand* /iudealand/ – "Judea" (Köbler 1989: 317; Snædal 1998: 514; Streitberg 2000: 165; Tollenaere & Jones 1976: 103);

6) *iudaiwiskon* /iudewiskon/ – inf. of the weak verb II "live like a Jew" (Köbler 1989: 317; Snædal 1998: 516; Streitberg 2000: 355; Tollenaere & Jones 1976: 104);

7) *Iudaius* /iudeus/ – "Jew" (Köbler 1989: 317; Snædal 1998: 516; Streitberg 2000: 81; Tollenaere & Jones 1976: 104), etc.

Two three-phoneme initial vocalic clusters were identified. They are represented by single usages in the Gothic Bible, e.g.:

1) /iai/ in *Iaeirus* – "Jairus", a ruler of the synagogue's name (Snædal 1998: 415; Streitberg 2000: 125; Tollenaere & Jones 1976: 85);

2) /ioa/ in Ioanan – "Johanan", the Old Testament proper name (Snædal 1998: 473; Streitberg 2000: 451; Tollenaere & Jones 1976: 98).

#### 4.3 Gothic two-phoneme final vocalic clusters

Besides, there are 37 (0.39% of the total number of word-forms under analysis) Gothic word-forms with the frequency of 801 (1.19% of the total word-form frequency) usages containing 19 two-phoneme vocalic clusters in the final position:

- 1) /iu/, e.g., Goth. *niu* – adv. "do not, did not, shall not, will not, have not" (Köbler 1989: 413; Snædal 1998: 775; Streitberg 2000: 219; Tollenaere & Jones 1976: 141);
- 2) /ui/, e.g., Goth. *puei* – relat. pron., nom., sing., "you who" (Köbler 1989: 557; Snædal 1998: 1089; Streitberg 2000: 245; Tollenaere & Jones 1976: 189);
- 3) /oe/, e.g., Goth. *ailoe* – interj. "my God" (Köbler 1989: 18; Snædal 1998: 27; Streitberg 2000: 221; Tollenaere & Jones 1976: 9);
- 4) /au/, e.g., Goth. *jau* – adv. "it may not be so that, is it so that, whether, regarding whether" (Köbler 1989: 333; Snædal 1998: 588; Streitberg 2000: 111; Tollenaere & Jones 1976: 116);
- 5) /ua/, e.g., Goth. *Iesua* – noun., dat., sing., masc. of *Iesus* "Jesus" (Snædal 1998: 420; Streitberg 2000: 59; Tollenaere & Jones 1976: 86);
- 6) /ii/, e.g., Goth. *pizeiei* – relat. pron., gen., pl., masc. of *saei* "the one who, the one which, which" (Köbler 1989: 458; Snædal 1998: 879; Streitberg 2000: 377; Tollenaere & Jones 1976: 187);
- 7) /io/, e.g., Goth. *gazaufwlakio* – noun., dat., sing., masc. of *gazaufwlakio* "treasury, treasure room" (Köbler 1989: 237; Snædal 1998: 313; Streitberg 2000: 43; Tollenaere & Jones 1976: 71);
- 8) /ia/, e.g., Goth. *praitoria* – noun., acc., sing., neut./fem. "praetorium, palace" (Köbler 1989: 420-421; Snædal 1998: 789; Streitberg 2000: 81; Tollenaere & Jones 1976: 143);
- 9) /ai/, e.g., Goth. *swaei* – conj. "so that, such that, so as to, therefore" (Köbler 1989: 510; Snædal 1998: 993; Streitberg 2000: 171; Tollenaere & Jones 1976: 167);

- 10) /ei/, e.g., Goth. *pizeei* – relat. pron., gen., pl., masc. and neut. of *saei* "the one who, the one which, which" (Köbler 1989: 458; Snædal 1998: 879; Streitberg 2000: 377; Tollenaere & Jones 1976: 187);
- 11) /oi/, e.g., Goth. *þaproei* – adv. "from the place that, from which place" (Köbler 1989: 543; Snædal 1998: 1063; Streitberg 2000: 377; Tollenaere & Jones 1976: 183);
- 12) /áii/, e.g., Goth. *þaiei* – relat. pron., nom., pl., masc. of *saei* "the one who, the one which, which" (Köbler 1989: 455; Snædal 1998: 879; Streitberg 2000: 177; Tollenaere & Jones 1976: 174);
- 13) /íai/, e.g., Goth. *siai* – 3<sup>rd</sup> pers. sing., pres. opt. of the irreg. verb *wisan* "be, exist, have existence, be present" (Köbler 1989: 638; Snædal 1998: 1219; Streitberg 2000: 391; Tollenaere & Jones 1976: 157);
- 14) /áua/, e.g., Goth. *gatraua* – 1<sup>st</sup> pers. sing., pres. ind. of the weak verb III *gatrauan* "trust, have confidence, be confident of, be convinced of, put into the trust of" (Köbler 1989: 228; Snædal 1998: 1031; Streitberg 2000: 229; Tollenaere & Jones 1976: 69);
- 15) /iáu/, e.g., Goth. *siau* – 1<sup>st</sup> pers. sing., pres. opt. of the irreg. verb *wisan* "be, exist, have existence, be present" (Köbler 1989: 640; Snædal 1998: 1219; Streitberg 2000: 131; Tollenaere & Jones 1976: 157);
- 16) /áiu/, e.g., Goth. *habaiu* – 3<sup>rd</sup> pers. sing., pres. opt. of the weak. Verb III *haban* followed by interrog. enclit. partic. *-u* "have, possess, have hold of, take hold of, have at one's disposal, consider" (Köbler 1989: 251; Snædal 1998: 344; Streitberg 2000: 141; Tollenaere & Jones 1976: 75);
- 17) /áuáu/, e.g., Goth. *gatrauau* – 1<sup>st</sup> pers. sing., pres. opt. of the weak verb III *gatrauan* "trust, have confidence, be confident of, be convinced of, put into the trust of" (Köbler 1989: 228; Snædal 1998: 1031; Streitberg 2000: 315; Tollenaere & Jones 1976: 69);
- 18) /áuái/, e.g., Goth. *bauai* – 3<sup>rd</sup> pers. sing., pres. opt. of the weak verb III *bauan* "dwell, inhabit, live" (Köbler 1989: 83; Snædal 1998: 133; Streitberg 2000: 389; Tollenaere & Jones 1976: 30);
- 19) /áui/, e.g., Goth. Goth. *táu* – noun., nom. and acc., sing., neut. "work, act, activity, deed, product, effect" (Köbler 1989: 523; Snædal 1998: 1017; Streitberg 2000: 335; Tollenaere & Jones 1976: 171).

There is only one three-phoneme final vocalic cluster /iáiu/ represented by a single word-form: Goth. *siaiu* – 3<sup>rd</sup> pers. sing., pres. opt. of the irreg. verb *wisan* followed by interrog. enclit. partic. *-u* "be, exist, have existence, be present" (Snædal 1998: 1220; Streitberg 2000: 141; Tollenaere & Jones 1976: 157).

The analysis of the initial vocalic clusters shows immediate phonological constraints that prevent their phonemic patterning, particularly in the word-forms of the Germanic origin. Instances of clusters in the Biblical proper names are transcriptions from the Greek original. Final vocalic clusters are conditioned mainly by the morphological system (noun, pronoun, and verbal paradigms). Vocalic patterning in Gothic demonstrates much lower combinability as compared to that of consonantal.

#### *4.4 Gothic consonantal clusters*

This section of the present study proposes a phonological analysis of initial (33 two-phoneme, two three-phoneme) and final (75 two-phoneme, 48 three-phoneme, and seven four-phoneme) consonantal clusters, based on a sample of 11988 (5038 two-phoneme initial, 5571 two-phoneme final, 20 three-phoneme initial, 1315 three-phoneme final, and 44 four-phoneme final) textual occurrences in the Gothic Bible and minor Gothic manuscripts.

##### *4.4.1 Gothic two-phoneme initial consonantal clusters*

Having analyzed all the corpus of the Gothic word-forms, 33 out of 361 (9.14% of realization) theoretically possible two-phoneme initial consonantal clusters were identified. Their frequency in the Gothic manuscripts is 5038 usages.

The most frequent two-phoneme consonantal clusters are:

1) /fr/ with 1634 usages, e.g., Goth. *fram* – prep. "forwards, forth, from, at, by, near, before, about" (Köbler 1989: 161-162; Snædal 1998: 244-247; Streitberg 2000: 7; Tollenaere & Jones 1976: 50-51);



2) /sw/ with 948 usages, e.g., Goth. *swe* – adv., conj. "as, just as, in like manner as, as if, like" (Köbler 1989: 513-514; Snædal 1998: 999-1002; Streitberg 2000: 211; Tollenaere & Jones 1976: 168-169);

3) /hr/ with 494 usages, e.g., Goth. *hropjan* – inf. of the weak verb I "cry out, shout, call out" (Köbler 1989: 275; Snædal 1998: 391; Streitberg 2000: 205; Tollenaere & Jones 1976: 81).

The least frequent two-phoneme consonantal clusters are:

1) /gl/ with 4 usages, e.g., Goth. *glaggwuba* – adv. "meticulously, with attention to detail, diligently" (Köbler 1989: 240; Snædal 1998: 323; Streitberg 2000: 85; Tollenaere & Jones 1976: 72);

2) /bn/ with a single usage, e.g., *bnauandans* – Part. I, nom., pl., masc. of the strong reduplicative verb *bnauan* "rub" (Köbler 1989: 100-101; Snædal 1998: 155; Streitberg 2000: 109; Tollenaere & Jones 1976: 35);

3) /k<sup>w</sup>r/ also with a single usage, e.g., *qrammipa* – noun, acc., sing., fem. of *qrammipa* "moisture" (Köbler 1989: 431; Snædal 1998: 820; Streitberg 2000: 121; Tollenaere & Jones 1976: 149).

The number of the consonant phonemes is 19. All of them except /j/ can occupy the first position in the initial cluster. The theoretically possible number of two-phoneme initial consonantal clusters is  $18^2 = 324$ . Of these only 33 occur (10.19% of realization). The second position in the cluster is occupied by 17 consonant phonemes; excluded from the sequences are the phonemes /h<sup>w</sup>/ and /j/. The most typical and frequent clusters are those containing sonorant phonemes /w/, /r/, /l/, /m/, and /n/ in the second position; they are all heterogeneous in regard to the presence in their phonemic structure modal positive kineme of obstruence. Their frequency in the corpus is 4360 (86.5% of the total occurrences).

The next clusters in the decreasing frequency order are those combined with the phoneme /s/ as the first member of the sequence, and obstruent plosives /p/, /t/, and /k/

– as the second. These three clusters are homogeneous in obstruency. They are found in 672 word-forms (13.33% of the total occurrences). Homogeneous cluster /ps/ and heterogeneous /bn/ demonstrate the least frequency: six cases of realization out of 5038.

#### 4.4.1 Gothic two-phoneme final consonantal clusters

There are 75 two-phoneme final consonantal clusters, which is 20.8% of the theoretically possible number of combinations. Their frequency in the Gothic corpus of word-forms is 5571 usages. The number of two-phoneme consonantal clusters in the final position is more than twice as high as the number of two-phoneme consonantal clusters in the initial position.

The most frequent two-phoneme final consonantal clusters are:

- 1) /ns/ with 2134 usages, e.g., Goth. *qumans* – Part. II, nom., sing., masc. of the strong verb IV *qiman* "come, arrive, get to" (Köbler 1989: 424-425; Snædal 1998: 793-798; Streitberg 2000: 115; Tollenaere & Jones 1976: 149);
- 2) /st/ with the frequency of 788 usages, e.g., Goth. *ist* – 3<sup>rd</sup> pers. sing., pres. ind. of the irreg. verb *wisan* "be, exist, have existence, be present" (Köbler 1989: 639-644; Snædal 1998: 1206-1214; Streitberg 2000: 123; Tollenaere & Jones 1976: 100-102);
- 3) /nd/ with the frequency of 669 usages, e.g., *letand* – 3<sup>rd</sup> pers. pl., pres. ind. of the strong reduplicative verb VII *letan* "let alone, leave, leave alone, let out, let forth" (Köbler 1989: 359; Snædal 1998: 55; Streitberg 2000: 159; Tollenaere & Jones 1976: 122).

Two-phoneme final consonantal clusters with the lowest frequency are represented by:

- 1) /h<sup>w</sup>t/ with the frequency of two usages, e.g., Goth. *sahut* – 2<sup>nd</sup> pers. sing., pres. ind. of the strong verb V *saihvan* "see, look, observe, view" (Köbler 1989: 461-462; Snædal 1998: 898-900; Streitberg 2000: 137; Tollenaere & Jones 1976: 154);
- 2) /mz/ with a single usage, e.g., Goth. *mimz* – noun, acc., sing., neut. of *mimz* "meat, edible flesh" (Köbler 1989: 388; Snædal 1998: 721; Streitberg 2000: 259; Tollenaere & Jones 1976: 132);

3) /lf/ also with a single usage, e.g., Goth. *wulf* – noun, acc., sing., masc. of *wulfs* "wolf" (Köbler 1989: 652; Snædal 1998: 1241-1242; Streitberg 2000: 51; Tollenaere & Jones 1976: 215).

All consonant phonemes except /j/, /h<sup>w</sup>/, and /j/ can occupy the first position in the final cluster. The theoretically possible number of two-phoneme initial consonantal clusters is  $16^2 = 256$ . Of these only 75 were attested (29.3% of realization). The phonemes /k<sup>w</sup>/, /w/, and /j/ are excluded by the Gothic language system from their role as the second elements of the final clusters.

Table 4. Combinability of the Gothic consonant phonemes in the two-phoneme initial and final clusters

	P H O N E M E S																		
	p	t	k	k <sup>w</sup>	f	θ	s	h/x	h <sup>w</sup>	b/v	d/ð	z	g/γ	m	n	w	r	l	j
P H O N E M E S	p						ps										pr	pl	
	t		tt				ts									tw	tr	tl	
	k						ks								kn		kr	kl	
	k <sup>w</sup>																k <sup>w</sup> r		
	f		ft				fs										fr	fl	
	θ						θs									θw	θr	θl	
	s	sp	st	sk			ss						sm	sn	sw		sl		
	h/x		ht				hs								hn	hw	hr	hl	
	h <sup>w</sup>		h <sup>w</sup> t																
	b/v						bs								bn		br	bl	
	d/ð						ds									dw	dr		
	z										zd				zn				
	g/γ		gt				gs				gd			gm	gn		gr	gl	
	m	mp	mt		mf		ms			mb		mz		mm					
	n		nt	nk	nk <sup>w</sup>	nθ	ns				nd	nz	ng		nn				
	w																wr	wl	
	r	rp	rt		rf	rθ	rs	rh		rb	rd		rg	rm	rn				
	l	lp	lt	lk	lf	lθ	ls	lh			ld		lg	lm				ll	
	j																		

Table 4 shows combinability of phonemes in two-phoneme initial and final consonantal clusters. Phonemes in the left side vertical column are the first members of the clusters, phonemes in the upright horizontal row are the second elements of the clusters. The attested consonantal clusters are realized at the intersection of phonemes. Clusters highlighted in blue are found only in the initial position; those in green are

final, while clusters in red are realized both initially and finally. Empty spaces mean exclusions on the phonemic combinability.

In Gothic, as shown in the Table 4, there are 16 two-phoneme consonantal clusters (highlighted in red), which are implemented both in the initial and final word-form positions. These combinations fall into two groups, based on their models. The first group of sequences /st, sk, ps/ contains two obstruent phonemes. It is noteworthy that the second and most numerous group of clusters /sn, fl, kn, kr, kl, θr, hn, hl, bn, br, dr, gl, gr/ is that of an obstruent combined with a sonorant phoneme. Two-phoneme final sequences of consonants (highlighted in green) display fewer constraints and exclusions on phonemic patterning as compared to the initial combinations (highlighted in blue), in which their realizations depend on and are regulated by the increase of sonority before the nucleus of the syllable represented by a vowel phoneme.

See the full list of the two-phoneme initial and final consonantal clusters in the frequency decreasing order in Tables 5 and 6 (Appendices A and B).

#### 4.4.3 Gothic three-phoneme initial consonantal clusters

Two three-phoneme initial consonantal clusters with a strict sequence of phonemes (/s + p / t + r/) were attested:

- 1) /spr/ with the frequency of 17 usages, e.g., Goth. *sprauto* – adv., conj. "quickly, speedily, promptly, fast, rapidly, without delay" (Köbler 1989: 496; Snædal 1998: 966-967; Streitberg 2000: 57; Tollenaere & Jones 1976: 168-169);
- 2) /str/ with the frequency of 3 usages, e.g., *strawidedun* – 3<sup>rd</sup> pers. pl., pret. ind. of the weak verb I *straujan* "strew, scatter, spread by throwing" (Köbler 1989: 503; Snædal 1998: 979; Streitberg 2000: 205; Tollenaere & Jones 1976: 165).

#### 4.4.4 Gothic three-phoneme final consonantal clusters

Analysis of three-phoneme consonantal clusters is complicated by a number of theoretical differences on this point. The inventory of three-phoneme consonantal

clusters, compiled by Schulze (1966), includes 36 consonantal clusters, of which 21 clusters are formed by combining two-phoneme clusters with consonant phonemes /s/, /t/, which are word-forming suffixes. Calabrese (1994) holds the same opinion. Thus, the realization of the three-phoneme final consonantal clusters in the word-forms of Gothic is supported by the morphological system of the language.

In the corpus of word-forms, there are 48 three-phoneme final consonantal clusters, which is 0.7% of the 6859 theoretically possible combinations. The frequency of these clusters is 1315 usages. The most frequent three-phoneme final consonantal clusters are:

1) /nds/, which is registered in 969 word-form usages, e.g., Goth. *ungalauþjands* – Part. I, nom., sing., masc. "unbelieving, uncompliant, disobedient" (Köbler 1989: 576; Snædal 1998: 1110-1111; Streitberg 2000: 269; Tollenaere & Jones 1976: 194);

2) /hts/ with the frequency of 73 usages, e.g., Goth. *waihts* – noun, nom. and gen., sing., acc., pl. of *waihts* "thing, anything, something, entity, objective, matter" (Köbler 1989: 611-612; Snædal 1998: 1147-1148; Streitberg 2000: 47; Tollenaere & Jones 1976: 205);

3) /nst/ with the frequency of 23 usages, e.g., Goth. *anst* – noun, acc., sing., fem. of *ansts* "meat, beneficence, graciousness, grace, joy, thanks, gift" (Köbler 1989: 53-54; Snædal 1998: 86-87; Streitberg 2000: 87; Tollenaere & Jones 1976: 22).

The least frequent three-phoneme final consonantal clusters are represented by single word-forms:

1) /rθr/ as in Goth. *maurþr* – noun, acc., sing., neut. of *maurþr* "murder, homicide" (Köbler 1989: 381; Snædal 1998: 704; Streitberg 2000: 219; Tollenaere & Jones 1976: 129);

2) /þms/ as in Goth. *maiþms* – noun, acc., sing., neut. of *maurþr* "votive treasure, gift" (Köbler 1989: 371; Snædal 1998: 682; Streitberg 2000: 187; Tollenaere & Jones 1976: 125);

3) /mbs/ as in Goth. *dumbs* – adj, nom., sing., masc. "voiceless, mute" (Köbler 1989: 129; Snædal 1998: 203; Streitberg 2000: 87; Tollenaere & Jones 1976: 42).

According to their structure, all the 48 three-phoneme final consonantal clusters are distributed among five groups. The most frequent clusters are described by such sequence of phonemes as a sonorant followed by two successive obstruent phonemes. This group is represented by 23 consonantal sequences. Their textual occurrence is the highest among the other four groups (1141 usages, which is 86.77% of the three-phoneme final clusters total frequency). The second group of clusters is represented by seven consonantal sequences modelled in the following order: first, second, and third positions are all occupied by obstruent phonemes. These consonantal sequences are attested in 115 word-forms (8.75% of the total three phoneme final clusters frequency). Another group of eight clusters is built according to such model, as an obstruent + a sonorant + an obstruent. Its textual frequency is 27 occurrences (2.05%). The fourth group of seven clusters with the frequency of 17 usages (1.29%) are constructed in the following order: a sonorant + an obstruent + a sonorant. The rarest group of the three-phoneme consonant sequences is an obstruent phoneme followed by another obstruent + a sonorant. Its frequency is 15 usages in the corpus under analysis.

See the full list of the three-phoneme final consonantal clusters in the frequency decreasing order in Table 7 (Appendix C).

#### 4.4.5 Gothic four-phoneme final consonantal clusters

In the Gothic corpus under analysis, seven four-phoneme final consonantal clusters with frequency of 44 usages were attested. These clusters are:

1) /nsts/ as in Goth. *ansts* – noun, nom., sing., fem. "meat, beneficence, graciousness, grace, joy, thanks, gift" (Köbler 1989: 53-54; Snædal 1998: 86-87; Streitberg 2000: 87; Tollenaere & Jones 1976: 22);

- 2) /*mfsl*/ as in Goth. *swumfsl* – noun, acc., sing., neut. "pool, pond, place of swimming" (Köbler 1989: 519; Snædal 1998: 1010; Streitberg 2000: 49; Tollenaere & Jones 1976: 170);
- 3) /*strs*/ as in Goth. *swistrs* – noun, gen., sing., fem. of *swistar* "sister" (Köbler 1989: 518; Snædal 1998: 1009; Streitberg 2000: 55; Tollenaere & Jones 1976: 170);
- 4) /*rkns*/ as in Goth. *airkns* – adj, nom., sing., masc. "holy, genuine, pure, sincere, unadulterated" (Köbler 1989: 22; Streitberg 2000: 419; Tollenaere & Jones 1976: 11);
- 5) /*rfts*/ as in Goth. *paurfts* – noun, nom., sing., fem. "need, necessity, requisiteness" (Köbler 1989: 544; Snædal 1998: 1065; Streitberg 2000: 157; Tollenaere & Jones 1976: 183);
- 6) /*rhts*/ as in Goth. *frawaurhts* – noun, nom., sing., fem. "evil-doing, offence, transgression" (Köbler 1989: 168; Snædal 1998: 51; Streitberg 2000: 157; Tollenaere & Jones 1976: 53);
- 7) /*hsns*/ as in Goth. *garehsns* – noun, nom., sing., fem. "design, plan, programme, determination, fixed time" (Köbler 1989: 216; Snædal 1998: 304; Streitberg 2000: 460; Tollenaere & Jones 1976: 66).

As seen from the above-drawn instances, four models of the four-phoneme consonantal clusters are realized according to the following models:

- 1) a sonorant + an obstruent + an obstruent /t/ + an obstruent /s/ – /*nsts*/, /*rfts*/, and /*hts*/;
- 2) an obstruent + an obstruent + a sonorant + an obstruent /s/ – /*strs*/ and /*hsns*/;
- 3) a sonorant /r/ + an obstruent /k/ + a sonorant /n/ + an obstruent /s/;
- 4) a sonorant /m/ + an obstruent /f/ + an obstruent /s/ + a sonorant /l/.

## 5. Discussion and conclusions

The study is focused on the search for such characteristics that would help corroborate the hypothesis that the corpus of the Gothic word-forms is not just a multitude of word-forms the language system is comprised of, but an ordered set of elements that are related in certain ways, not only at the grammatical and semantic levels, but also at the

phonemic one. The word-form as a linguistic unit is not a simple sequence of phonemes, it is also structured at the phonemic level.

Phonemic patterning of the word-forms was studied in the aspect of its structural characteristics, calculating the average word-form length in phonemes, setting the oscillation limits of average and absolute lengths, establishing the margin of error by statistical analysis. Combinability of the two classes of phonemes (vowels and consonants) within a word-form was undertaken by applying the method of distributional analysis at the segmental (phonemic) level.

The study of the word-form length in phonemes has shown that the Gothic language macrosystem imposes restrictions on its most general structural characteristics – the number of phonemes that constitute a word-form. There are no word-forms longer than 19 phonemes in the Gothic language; neither has any 18-phoneme long word-form been registered. All registered word-forms are grouped into 18 classes according to the number of phonemes in their structure. Each class has a different number of word-forms, the largest being the seven-phoneme class. Close in number to this class are the six- and eight-phoneme-long word-forms. With the increase of length in phonemes, the number of word-forms in a certain class decreases. The longer in phonemes the word-forms are, the less convenient they are for the speakers of the language.

The analysis of the general patterns in the phonemic structure of word-forms has shown that inside the word-forms of each length, the Gothic language imposes its restrictions and limitations on the combinability of the two classes of phonemes – vowels and consonants, presented in the most general form as V and C. A comparison of the theoretically possible and instantiated canonical forms shows that the Gothic language exploits its potential by less than one percent, while the number of theoretically possible canonical forms used by the language decreases with their increasing length. The inquiry into the canonical forms has shown that the Gothic language does not allow



vocalic clusters longer than three phonemes and consonantal clusters longer than four phonemes successively in word-forms of any length.

Some of the canonical forms cannot be word-forms according to the laws of any Indo-European language, being only clusters of consonant phonemes. Others could theoretically be word-forms, but they are not found in the Gothic manuscripts. Among the unrealized constructional patterns, which, according to the laws of the Gothic language, could be word-forms, since the combination of vowel and consonant phonemes in them does not exceed the norm established in the language, they, nevertheless, do not model a single word. Obviously, the restrictions imposed by the Gothic language on the combinability of phonemes belonging to the same class in a word-form must be considered in close connection with its length: the shorter the canonical forms are, the shorter the same class phoneme clusters are, even though such clusters are permitted by the laws of the language.

The analysis of the modelling power of canonical forms measured by the number of word-forms and their textual frequency has shown that the presence of three-phoneme vocalic or four-phoneme consonantal clusters, which are allowed by the language, sharply reduces the modelling power of a canonical form. The inventory of 71 (out of 656) most frequent canonical forms with cumulative modelling power of 90% has been determined as the centre of the system. The remaining inventory of 585 canonical forms, whose cumulative modelling power is 10%, represents the system's periphery.

The distributional analysis of the initial vocalic clusters has revealed one two-phoneme sequence of /i/ +/u/ in the word-forms of the Germanic origin. The other seven two-phoneme and two three-phoneme initial clusters have been found in proper names of the Biblical personages, geographical names, and names of nationalities, which were transcribed from the Old Greek New Testament original, or via Old Greek from the Hebrew Old Testament, where such combinations of vowel phonemes are permitted by the laws of the respective languages. The examination of the two-phoneme and

three-phoneme final vocalic sequences has shown that their combinability is conditioned mainly by the noun, pronoun, and verbal paradigms.

The study of consonantal clusters has revealed a significantly larger inventory of construction patterns (33 initial two-phoneme, two three-phoneme combinations), whereas 75 two-phoneme, 48 three-phoneme, and seven four-phoneme sequences have been attested in the end of the word-forms. The most typical and frequent two-phoneme clusters are those containing sonorant phonemes as the first members of the sequences, the second position is occupied by any phoneme containing a positive modal kineme of obstruence. Next in the frequency decreasing order are the clusters with obstruent phonemes in the first position. The second position is occupied by obstruent phonemes or sonorants. Among the least frequent two-phoneme final consonantal clusters are those consisting of two sonorant phonemes, an obstruent phoneme followed by a sonorant, or by another obstruent.

The investigation of the two three-phoneme initial consonantal clusters has identified a strict order of their implementation: the first member of the sequence is the phoneme /s/, the second position is occupied by the obstruent phonemes /p/ or /t/ followed by the sonorant phoneme /r/. The treatment of the final consonantal clusters has shown that they are patterned according to five models. The most frequent clusters are represented by such sequence of phonemes as a sonorant followed by two successive obstruent phonemes. The next group of clusters in the frequency decreasing order is modelled by the obstruent phonemes realized successively. Another group of consonantal combinations is instantiated by the sequence of an obstruent + a sonorant + an obstruent. Consonantal sequences belonging to the fourth group are constructed in the following order: a sonorant + an obstruent + a sonorant. The rarest group of the three-phoneme consonantal sequences is an obstruent followed by another obstruent + a sonorant phoneme.

The application of statistical analysis methods has revealed such features of the phonemic structure of the Gothic word-form, as its length in phonemes, the position of phonemes in the word-form, and the position of classes of phonemes in the canonical forms. Statistical analysis confirmed the reliability of the obtained results: the relative margin of error does not exceed 0.5%, and the accuracy of statistical characteristics is 99.5%.

The proposed methodology, results, and conclusions of the study can be used in further theoretical developments not only in phonology, but also in the study of typological features of the cognate and genetically distant languages belonging to different grammatical types.

### **Abbreviations**

acc. – accusative

adj. – adjective

adv. – adverb

CF – canonical form

conj. – conjunction

demonstr. – demonstrative

fem. – feminine

gen. – genitive

dat. – dative

enclit. – enclitic

Goth. – Gothic

ind. – indicative

inf. – infinitive

interj. – interjection

irreg. – irregular

masc. – masculine

neut. – neuter

nom. – nominative

Part. I – present participle

Part II – past participle

partic. – particle

pers. – person

pl. – plural

prep. – preposition

pres. – presence

pret.-pres. – preterite-present

pret. – preterite

pron. – pronoun

relat. – relative

sing. – singular

supplet. – suppletive

## References

Agud Aparacio, A. & Fernandez Alvarez, M.P. (1988). *Manual de la lengua gótica*. Salamanca: Ediciones Universidad de Salamanca.

Akidah, M.A. (2013). *Lexical phonology and morphology of modern standard Arabic*. Saarbrücken: Lambert Academic Publishing.

Antonsen, E.H. (1972). The Proto-Germanic syllabics (vowels). In *Toward a grammar of Proto-Germanic*. Coetsem, F. van & Kufner, H.L. (eds.). Tübingen: Max Niemeyer Verlag, p. 117-140.

Barrack, Ch.M. (1998). *Siever's law in Germanic*. New York: Peter Lang.

Baudouin de Courtenay, J. (1972a). An attempt at a theory of phonetic alternations. In Baudouin de Courtenay, J. *The beginnings of structural linguistics: Anthology*. Stankiewicz, E. (ed.). Bloomington – London: Indiana University Press, p. 144-212.

Baudouin de Courtenay, J. (1972b). An outline of the history of the Polish language. In Baudouin de Courtenay, J. *The beginnings of structural linguistics: Anthology*. Stankiewicz, E. (ed.). Bloomington – London: Indiana University Press, p. 323-355.

- Baudouin de Courtenay, J. (1972c). The difference between phonetics and psychophonetics. In Baudouin de Courtenay, J. *The beginnings of structural linguistics: Anthology*. Stankiewicz, E. (ed.). Bloomington – London: Indiana University Press, p. 278-283.
- Bennett, W.Y. (1980). *An introduction to the Gothic language*. New York: The Modern Language Association of America.
- Benveniste, É. (1966). *Problèmes de linguistique generale*. Paris: Gallimard.
- Braune, W. (1883). *Gothic grammar: With selections for reading and a glossary*. London: Sampson Low, Marston, Searle & Rivington.
- Calabrese, A. (1994). Sievers' Law in Gothic: A synchronic analysis with some notes on its diachronic development. In *The linguistic review*, 11, p. 149-194.
- Chomsky, N. & Halle, M. (1991). *The sound pattern of English*. Cambridge – London: The MIT Press.
- Daneš, F. (1966). The relation of centre and periphery as a language universal. In *Travaux linguistiques de Prague. 2. Les problèmes du centre et la périphérie du système de la langue*. Vachek, J. (ed). Tuscaloosa: University of Alabama Press, p. 9-21.
- Ebbinghaus, E.A. (1971). Gotica II. In *General linguistics*, 11, 2, p. 99-103.
- Feist, Z. (1939). *Vergleichendes Wörterbuch der gotischen Sprache*. Leiden: E.J. Brill.
- Fischer-Jørgensen, E. (1995). *Trends in phonological theory until 1975: A historical introduction*. Copenhagen: C.A. Reitzel.
- Fulk, R.D. (2018). *A comparative grammar of the early Germanic languages*. Amsterdam – Philadelphia: John Benjamins Publishing Company.
- Gamkrelidze, T.V. & Ivanov, Viach.Vs. (1995). *Indo-European and the Indo-Europeans: A reconstruction and historical analysis of a proto-language and a proto-culture*. Berlin – New York: Mouton de Gruyter. Vol. I, II.
- Greenberg, J.H. (1990). Some areal characteristics of African languages. In *On language: Selected writings of Joseph H. Greenberg*. Denning, K. & Kemmer, S. (eds.). Stanford: Stanford University Press, p. 484-501.

Grucza, F. (1970). *Sprachliche Diakrise im Bereich der Ausdrucksebene des Deutschen: Beiträge zur allgemeinen Sprachtheorie*. Poznań: Państwowe wydawnictwo naukowe.

Gukhman, M.M. (1958). *The Gothic language*. Moscow: Izdatel'stvo literatury na inostrannykh yazykakh. / Gukhman M.M. *Gotskij yazyk*. Moskva: Izdatel'stvo literatury na inostrannykh yazykakh. / Гухман М.М. *Готский язык*. Москва: Издательство литературы на иностранных языках.

Jakobson, R. (1971). On the identification of phonemic entities. In Jakobson, R. *Selected writings*. Vol. I. Phonological studies. 2<sup>nd</sup> ed., exp. The Hague – Paris: Mouton, p. 418-425.

Jakobson, R. & Halle, M. (1971). Phonology and phonetics. In Jakobson, R. *Selected writings*. Vol. I. Phonological studies. 2<sup>nd</sup> ed., exp. The Hague – Paris: Mouton, p. 464-504.

Jellinek, M.H. (1926). *Geschichte der gotischen Sprache*. Berlin – Leipzig: Walter de Gruyter & Co.

Joos, M. (1942). Statistical patterns in Gothic phonology. In *Language*, 18, 1, p. 33-38.

Köbler, G. (1989). *Gotisches Wörterbuch*. Leiden – New York – København – Köln: E.J. Brill.

Krámský, J. (1974). *The phoneme: Introduction to the history and theories of a concept*. München: Wilhelm Fink Verlag.

Marchand, J.W. (1973). *The sounds and phonemes of Wulfila's Gothic*. The Hague – Paris: Mouton.

Miller, D.G. (2019). *The Oxford Gothic grammar*. Oxford: Oxford University Press.

Moradi, H. & Chen, J. (2018). A contrastive analysis of Persian and English vowels and consonants. In *Lege Artis. Language yesterday, today, tomorrow*. The journal of University of SS Cyril and Methodius in Trnava. Warsaw: De Gruyter Open, III (2), December 2018, p. 105-131. DOI: 10.2478/lart-2018-0016

Mossé, F. (1969). *Manuel de la langue gotique: Grammaire, textes, glossaire*. 2<sup>nd</sup> ed. Paris: Aubier, Editions Montaigne.

- Moulton, W.G. (1972). The Proto-Germanic non-syllabics (consonants). In *Toward a grammar of Proto-Germanic*. Coetsem, F. van & Kufner, H.L. (eds.). Tübingen: Max Niemeyer Verlag, p. 141-173.
- Nielsen, H.F. (2013). The Gothic language of bishop Wulfila: Phonology, typology and purported purity. In *Acta bibliothecae r. Universitatis Upsaliensis, Vol. XLVIII, Opia 57. Wulfila 311-2011. International simposium, Uppsala University*, June 15-18, 2011, p. 179-196.
- Park, S. (2019). Construction patterns of the root morpheme in the Gothic language. In *Science and education: A new dimension. Philology*, 189, p. 46-49. / Park S. Modeli pobudovy korenevoyi morfemy u hots'kij movi. In *Science and education: A new dimension. Philology*, 189, s. 46-49. / Парк С. Моделі побудови кореневої морфеми у готській мові. In *Science and education: A new dimension. Philology*, 189, с. 46-49.
- Park, S. (2018a). Phonological architecture of word-forms in Gothic. In *Studia philologica*, 11, p. 4-10. / Park S. Fonolohichna arkhitektura slovoform u hots'kij movi. In *Studia philologica*, 11, s. 4-10. / Парк С. Фонологічна архітектура словоформ у готській мові. In *Studia philologica*, 11, с. 4-10.
- Park, S. (2018b). The system of phonemes and the system of phonemic oppositions of the Gothic language. In *Problems of semantics, pragmatics and cognitive linguistics*, 34, p. 38-51. / Park S. Systema fonem i systema fonematychnykh opozytsij hots'koyi movy. In *Problemy semantyky, prahmatyky ta kohnityvnoyi linhvistyky*, 34, s. 38-51. / Парк С. Система фонем і система фонематичних опозицій готської мови. In *Проблеми семантики, прагматики та когнітивної лінгвістики*, 34, с. 38-51.
- Perebyjnis, V.S. (1970). *Qantitative and qualitative characteristics of the system of phonemes of the modern Ukrainian literary language*. Kyiv: Naukova dumka. / Perebyjnis V.S. *Kil'kisni ta yakisni kharakterystyky systemy fonem suchasnoyi ukrayins'koyi literaturnoyi movy*. Kyuiv: Naukova dumka. / Перебийніс В.С. *Кількісні та якісні характеристики системи фонем сучасної української літературної мови*. Київ: Наукова думка.

- Perebyjnis, V.I. (2002). *Statistical methods for linguists*. Vinnytsya: Nova knyha publishers. / Perebyjnis V.I. *Statystychni metody dlya lnhvistiv*. Vinnytsya: Nova knyha. / Перебийніс В.І. *Статистичні методи для лінгвістів*. Вінниця: Нова книга.
- Pike, K.L. (1947). Grammatical prerequisites to phonemic analysis. In *Word*, 3, p. 155-172.
- Plotkin, V. (2008). *The evolution of Germanic phonological systems: Proto-Germanic, Gothic, West Germanic, and Scandinavian*. Lewiston – Queenston – Lampeter: The Edwin Mellen Press.
- Plotkin, V.Y. (1978). The kinakeme as the ultimate unit of language. In *Kwartalnik neofilologiczny*, xxv (3), p. 275-288.
- Rauch, I. (2011). *The Gothic language: Grammar, genetic provenance and typology, readings*. 2<sup>nd</sup> ed. New York – Washington – Baltimore – Bern – Frankfurt am Main – Berlin – Brussels – Vienna – Oxford: Peter Lang.
- Rauch, I. (2017). Toward schwa in Gothic again and its melody. In *Sprachwissenschaft*, 42 (3), p. 231-245.
- Rousseau, A. (2016). *Gotica: Études sur la langue gotique*. Paris: Honoré Champion Éditeur.
- Schulze, W. (1966). Wortbrechung in den gotischen Handschriften. In *Kleine Schriften*, 2. Auflage. Göttingen: Vandenhoeck & Ruprecht, p. 483-496.
- Smirnitskij, A.I. (1954). The word revisited. (The problem of the word "identity"). In *Proceedings of the Institute of Linguistics of the USSR Academy of Sciences*. Moscow: The USSR Academy of Sciences Publishing. Vol. IV, p. 3-49. / Smirnitskij A.I. К вопросу о слове. (Проблема тождества слова"). In *Trudy Instituta yazykoznaniya AN SSSR*. T. IV, s. 3-49. / Смирницкий А.И. К вопросу о слове (Проблема "тождества слова"). In *Труды Института языкознания АН СССР*. Москва: АН СССР. Т. IV, с. 3-49.
- Snædal, M. (1998). *A concordance to Biblical Gothic II*. Reykjavík: Institute of linguistics, University of Iceland, University of Iceland Press.



- Snædal, M. (2013). Gothic letter (and phoneme) statistics. In *Studia linguistica Universitatis Iagellonicae Cracoviensis*, 130, p. 277-295.
- Snædal, M. (2009). Ostgermanische Morphologie. In *Proceedings of the International conference "Morphology and digitalisation"*, Berlin, July 7-8, 2006, Chatreššar, p. 147-167.
- Streitberg, W. (2000). *Die Gotische Bibel. Der gotische Text und seine griechische Vorlage. Mit Einleitung, Lesarten und Quellennachweisen sowie den kleineren Denkmälern als Anhang*. Mit einem Nachtrag von Piergiuseppe Scardigli. Band I. 7. Auflage. Heidelberg: Universitätsverlag C. Winter.
- Tollenaere, F. de & Jones, R.L. (1976). *Word-indices and word-lists to the Gothic Bible and minor fragments*. Leiden: E.J. Brill.
- Torsuyev, G.P. (1962). *Issues of the phonetic structure of the word*. Moscow: Academy of Sciences of the USSR Publishing. / Torsuyev G.P. *Voprosy foneticheskoy struktury slova*. Moskva: Izdatel'stvo AN SSSR. / Торсуев Г.П. *Вопросы фонетической структуры слова*. Москва: Издательство АН СССР.
- Trubetzkoy, N.S. (1969). *Principles of phonology*. Berkeley – Los Angeles: University of California Press.
- Ufimtseva, A.A. (1962). *Experience of studying lexicon as a system*. Moscow: Academy of Sciences of the USSR Publishing. / Ufimtseva A.A. *Opyt izucheniya leksiki kak sistemy*. Moskva: Izdatel'stvo AN SSSR. / Уфимцева А.А. *Опыт изучения лексики как системы*. Москва: Издательство АН СССР.
- Vachek, J. (1976a). On the integration of the peripheral elements into the system of language. In *Selected writings in English and general linguistics*. Prague: Academia, p. 77-90.
- Vachek, J. (1976b). Phonemes and phonological units. In *Selected writings in English and general linguistics*. Prague: Academia, p. 15-19.
- Vachek, J. (1966). *The linguistic school of Prague: An introduction to its theory and practice*. Bloomington – London: Indiana University Press.
- Vasko, R.V. (1997). *Linear syntagmatics of kinakemes in the consonantal clusters of Gothic*. Synopsis for the thesis for the candidate degree in philology. Speciality

10.02.04 – Germanic languages. Kyiv: Kyiv State Linguistic University. / Vas'ko R.V. *Linijna syntahmatyka kinakem u konsonantnykh hrupakh hots'koyi movy*. Avtoreferat dissertatsiyi na zdobuttya naukovoho stupenya kandydata filolohichnykh nauk. Spetsial'nist' 10.02.04 – hermans'ki movy. Kyiv: Kyivskiy derzhavnyy lnhvistychnyj universytet. / Васько Р.В. *Лінійна синтагматика кінакем у консонантних групах готської мови*. Автореферат дисертації на здобуття наукового ступеня кандидата філологічних наук. Спеціальність 10.02.04 – германські мови. Київ: Київський державний лінгвістичний університет.

Vasko, R.V. (2007). *The ultimate phonological units in the consonantal system of the Old Germanic languages*. Synopsis for the thesis for the doctor degree in philology. Speciality 10.02.04 – Germanic languages. Kyiv: Kyiv National Linguistic University. / Vas'ko R.V. *Pervynni fonolohichni odynytsi v systemi konsonantyzmu davn'ohermans'kykh mov*. Avtoreferat disertatsiyi na zdobuttya naukovoho stupenya doktora filolohichnykh nauk. Spetsial'nist' 10.02.04 – hermans'ki movy. Kyiv: Kyivskiy natsional'nyy lnhvistychnyj universytet. / Васько Р.В. *Первинні фонологічні одиниці в системі консонантизму давньогерманських мов*. Автореферат дисертації на здобуття наукового ступеня доктора філологічних наук. Спеціальність 10.02.04 – германські мови. Київ: Київський національний лінгвістичний університет.


Vennemann, Th. (1988). *Preference laws for syllable structure and the explanation of sound change*. Berlin – New York – Amsterdam: Mouton de Gruyter.

Voyles, J.B. (1992). *Early Germanic grammar*. San Diego – New York – Boston – London – Sydney – Tokyo – Toronto: Academic Press, Harcourt Brace Jovanovich Publishers.

Wright, J. (1910). *Grammar of the Gothic language and the Gospel of St. Mark, selections from the other Gospels and the second Epistle to Timothy, with notes and glossary*. Oxford: Clarendon Press.

Zadorozhnyj, B.M. (1960). *Comparative phonetics and morphology of the Gothic language*. Lviv: Lviv University Press. / Zadorozhnyj B.M. *Porivnyal'na fonetyka i morfolohiya hots'koyi movy*. L'viv: Vydavnytstvo L'vivskoho universytetu. /

Задорожний Б.М. *Порівняльна фонетика і морфологія готської мови*. Львів: Видавництво Львівського університету.

<b>Contact data</b> Seeun Park M.A. in Philology PhD Student Prof. G.G. Pocheptsov Department of Germanic and Finno-Ugric Philo- logy, Kyiv National Linguistic University, Velyka Vasylkivska Str. 73, Kyiv, 03150, Ukraine, e-mail: <a href="mailto:seiun0725@gmail.com">seiun0725@gmail.com</a>		<b>Fields of interest</b>  Historical linguistics, language typology, comparative grammar of Germanic languages, diachronic phonology and morphology, language universals.
--	--	--

## Résumé

This study of the phonemic patterning in the Gothic word-forms aims at quantitative, distributional, and statistical analysis of the phonological structure of the word-forms attested in Gothic manuscripts at the segmental level. It is based on the hypothesis that the systemic character of meaningful linguistic units is corroborated not only by grammatical and semantic criteria, but also by the phonological one. The word-form is not a simple sequence of phonemes, it is also structured at the phonemic level. Calculations have shown that the Gothic language, theoretically capable of building the inventory of more than 1000000 word-forms construction patterns, actually utilized only 656, or 0.06 per cent. Any vocalic or consonantal cluster would inevitably bring about a corresponding reduction in the modelling power of the canonical form. The analysis of canonical forms has shown that the Gothic language does not allow vocalic and consonantal clusters longer than three and four phonemes respectively. The study of the word-form length in phonemes has shown that the Gothic language imposes restrictions and exclusions on its most general structural characteristics – the number of phonemes that constitute a word-form. Each of the 18 classes that were grouped according to the length in phonemes, is characterized by a different number and textual

frequency, the largest being the class of the word-forms containing seven phonemes. With the increase of length in phonemes, the number and frequency of word-forms in a certain class decreases. Distributional analysis at the phonemic level has made it possible to establish all initial and final vocalic and consonantal clusters and their constructional models. The choice of phonemes and number of clusters in the word-form initial position is very strict, even rigid, while the final position demonstrates more freedom for the cluster's realization. Statistical analysis has confirmed the reliability of the obtained results.

**Key words:** word-form, phoneme, kineme, phonemic structure, distribution, canonical form.

Article was received by the editorial board 12.02.2020;

Reviewed 22.03.2020. and 25.03.2020.

Similarity Index 1%

## Appendix A

Table 5. Two-phoneme initial consonantal clusters in the frequency decreasing order

#	Cluster	Frequency	#	Cluster	Frequency	#	Cluster	Frequency
1.	/fr/	1634	12.	/pr/	76	23.	/hn/	12
2.	/sw/	948	13.	/tr/	75	24.	/θl/	11
3.	/hr/	494	14.	/gr/	54	25.	/wl/	11
4.	/sk/	332	15.	/sp/	53	26.	/kn/	10
5.	/st/	288	16.	/sl/	49	27.	/pl/	7
6.	/br/	284	17.	/wr/	47	28.	/fl/	6
7.	/hl/	143	18.	/sn/	39	29.	/ps/	5
8.	/dr/	110	19.	/θw/	37	30.	/kl/	4
9.	/tw/	89	20.	/dw/	14	31.	/gl/	4
10.	/θr/	88	21.	/sm/	13	32.	/bn/	1
11.	/bl/	87	22.	/kr/	12	33.	/k <sup>w</sup> r/	1

## Appendix B

Table 6. Two-phoneme final consonantal clusters in the frequency decreasing order

#	Cluster	Frequency	#	Cluster	Frequency	#	Cluster	Frequency
1.	/ns/	2134	26.	/nθ/	17	51.	/hn/	3
2.	/st/	788	27.	/zn/	16	52.	/nk/	3
3.	/nd/	669	28.	/ss/	16	53.	/lg/	3
4.	/rθ/	224	29.	/sn/	16	54.	/br/	3
5.	/θs/	216	30.	/lk/	16	55.	/tt/	3
6.	/rh/	208	31.	/rs/	14	56.	/lm/	2
7.	/rd/	143	32.	/lθ/	13	57.	/h <sup>w</sup> t/	2
8.	/ht/	134	33.	/θr/	12	58.	/rt/	2
9.	/nn/	112	34.	/rp/	12	59.	/mm/	2
10.	/ll/	100	35.	/kn/	11	60.	/gl/	2
11.	/gs/	83	36.	/fs/	10	61.	/nz/	1
12.	/ks/	54	37.	/mt/	9	62.	/lf/	1
13.	/ng/	49	38.	/gn/	9	63.	/rf/	1
14.	/ds/	44	39.	/kl/	8	64.	/rm/	1
15.	/lh/	42	40.	/gt/	8	65.	/mz/	1
16.	/ms/	40	41.	/nt/	7	66.	/mb/	1
17.	/ls/	37	42.	/ps/	7	67.	/mp/	1
18.	/ts/	32	43.	/rb/	6	68.	/kr/	1
19.	/rg/	32	44.	/zd/	5	69.	/fl/	1
20.	/lt/	30	45.	/gm/	5	70.	/hw/	1
21.	/rn/	28	46.	/bs/	5	71.	/hl/	1
22.	/ld/	25	47.	/lp/	4	72.	/bn/	1
23.	/hs/	25	48.	/sk/	4	73.	/dr/	1
24.	/ft/	23	49.	/gr/	4	74.	/gd/	1
25.	/mf/	21	50.	/nk <sup>w</sup> /	4	75.	/tl/	1

## Appendix C

Table 7. Three-phoneme final consonantal clusters in the frequency decreasing order

#	Cluster	Frequency	#	Cluster	Frequency	#	Cluster	Frequency
1.	/nds/	966	17.	/lds/	6	33.	/kns/	2
2.	/hts/	73	18.	/sns/	6	34.	/kls/	2
3.	/rθs/	24	19.	/fst/	6	35.	/nnt/	1
4.	/lks/	23	20.	/θrs/	6	36.	/ngk/	1
5.	/nst/	23	21.	/hsn/	5	37.	/ngk <sup>w</sup> /	1
6.	/rht/	21	22.	/nsl/	5	38.	/hst/	1
7.	/sts/	19	23.	/rst/	4	39.	/mbs/	1
8.	/rgs/	16	24.	/zds/	4	40.	/mbn/	1
9.	/rds/	14	25.	/lθs/	3	41.	/msl/	1
10.	/nθs/	14	26.	/rft/	3	42.	/mts/	1
11.	/fts/	9	27.	/tts/	3	43.	/lht/	1
12.	/str/	8	28.	/rhs/	2	44.	/lfs/	1
13.	/rts/	7	29.	/rms/	2	45.	/rbs/	1
14.	/gms/	7	30.	/htr/	2	46.	/rθr/	1
15.	/lls/	6	31.	/ngs/	2	47.	/θms/	1
16.	/lhs/	6	32.	/krs/	2	48.	/brs/	1