Preprints.org

Article

Time Evolution of the Symmetry of Alphabet Symbols and Its Quantification: Study in the Archeology of Symmetry

[Artem Gilevich](#) , [Mark Frenkel](#) , [Shraga Shoval](#) , [Edward Bormashenko](#) [*]

Article

Time Evolution of the Symmetry of Alphabet Symbols and Its Quantification: Study in the Archeology of Symmetry

Artem Gilevich [1], Mark Frenkel [1], Shraga Shoval [2] and Edward Bormashenko [2,*]

[1] Department of Chemical Engineering, Ariel University, Ariel, POB 3, 407000, Israel

[2] Department of Industrial Engineering and Management, Faculty of Engineering, Ariel University, P.O.B. 3, Ariel 407000, Israel

* Correspondence: edward@ariel.ac.il

Abstract: Time evolution of the symmetry of symbols constituting alphabets rooted in the Phoenician script was investigated. A diversity of quantitative measures of symmetry of graphemes appearing in Phoenician, Western Greek, Etruscan from Marsiliana, Archaic Etruscan, Neo-Etruscan, Euclidian Greek, Archaic and Classical Latin, and Proto-Hebrew scripts, constituting the Phoenician script family, were calculated. The same measures were calculated for the Hebrew/Ashurit and English scripts. The Shannon-like measures of symmetry were calculated. The Shannon diversity index was calculated. The Shannon diversity index increased with time in a monotonic way for the scripts emerging from the Phoenician one. The diversity of symmetry groups inherent for alphabets emerging from the Phoenician one grows with time. We also introduced the symmetry factor of the alphabet. The symmetry factor quantifies the averaged level of symmetrization of the alphabet and the possible parsimony of graphical information necessary for the drawing of the entire set of graphemes, constituting the alphabet. The symmetry factor is decreased with time for the alphabets rooted in the Phoenician one. This means that the average level of symmetrization of the studied alphabet is increased with time. The parsimony of graphical information necessary for writing graphemes is increased with time. The values of the symmetry factor calculated for the addressed scripts are close one to another, with the pronounced exception of the Hebrew/Ashurit script. Our study supplies the arguments for the point of view, according to which the modern Hebrew/Ashurit script does not emerge from the Phoenician one.

Keywords: Phoenician alphabet; symmetry; symmetry group; grapheme; time evolution; Shannon measure of symmetry; Shannon diversity index

1. Introduction

Depending on how you count, there are 6000-8000 distinct languages on earth. The number of known scripts is much smaller, again, depending on how we count, we recognize *ca*. one hundred scripts. The term "script" denotes for a set of written marks together with conventions for using them to record a particular language; thus e.g., English and Finnish use the same alphabet, but their 'scripts' are rather different – English spelling being highly irregular and Finnish extremely regular [1]. Language has physical forms to be studied. We listen to speech, see writing and signing, and feel Braille dots read by the fingers [2]. The forms can be decomposed into structured components: sentences, phrases, words, letters, sounds [2]. In our manuscript we propose the physical approach to the time evolution of symbols constituting alphabets, emerging from the Phoenician alphabet [3–5]. Namely, we addressed the time evolution of symmetry of marks, constituting the scripts, rooted in the Phoenician alphabet. For a purpose of comparison, the modern English script was analyzed. The relation of the Hebrew/Ashurit script to the Phoenician group is disputable and it is addressed in the manuscript.

"Nothing in biology makes sense except in the light of evolution" is a famous dictum of the biologist Theodosius Dobzhansky. This means that the explanation of the phenomenon necessarily implies its analysis in the dimension of time, i.e. understanding of its time evolution [6]. This

statement is definitely true also for linguistics; thus, our research, devoted to the evolution of alphabets, is at least partially, related to evolutionary linguistics. And in equal measure our research is related to the mathematical theory of symmetry [7].

Time evolution of scripts attracted an attention of investigators in the last decade. A little is known about the origin of scripts. Some argue that symbolic graphical representations evolved from earlier iconic representations. For example, the Assyrian symbolic writing system evolved from the iconic pictographic system of early Cuneiform via early Babylonian [8]. A similar observation can be made about the evolution of Chinese characters [8]. Time evolution of symbols constituting West African languages was reported [9]. The crucial question is: do we recognize some distinct tendency in the time evolution of scripts? In other words: what are the laws, governing time evolution of scripts, constituting alphabets? It was already suggested in XIX century that letters, symbols of alphabets evolve to their simplification, i.e. with time the symbols of alphabets converge to more simple graphical forms [10,11]. This hypothesis was generalized and developed as follows: "through repeated interactions, a system of signs will become compressed so that the same amount of information is expressed with less descriptive effort" [10]. In other words, a more accurate wording which is adequate for describing the time evolution of alphabets is "compression" and not "simplification" [10]. This idea was exemplified recently by analyzing of Chinese characters spanning more than 3.000 years of recorded history [12]. No consistent evidence of simplification through time was revealed [11]. Moreover, it was found that modern Chinese characters are higher in visual complexity than their earliest known counterparts [12].

The reasonable question is: how the "simplicity" or "information compression" may be mathematically quantified? We demonstrate in our research that this challenging task may be accomplished with the analysis of the Shannon measures of symmetry of the symbols constituting the alphabets, as applied to alphabets emerging from the Phoenician one [13]. We also demonstrate that the Shannon measure of symmetry and the Shannon diversity index are an adequate mathematical tools, enabling quantitative description of symbols [13]. We also introduce the symmetry factor of the alphabet, which quantifies the average symmetry of its symbols. Our research develops the approach suggested by Revesz who pointed to the importance of symmetry for the analysis of the time evolution of alphabets [14].

2. Methods

2.1. Scripts addressed on the investigation

Phoenician, Etruscan from Marsiliana, Archaic Etruscan, Neo-Etruscan [16], Archaic Western Greek, Euclidian Greek, Archaic and Classical Latin [19,20], Proto-Hebrew scripts, constituting the Phoenician script family, along with Hebrew/Ashurit script, which relation to the Phoenician group is disputable were analyzed [22]. Scripts and shapes of the symbols are summarized in Appendix and Supplementary Materials. Analysis of the symmetry of the Phoenician script family exploited the procedure, introduced by Revesz [14]. This procedure considered the vertical symmetry of symbols [14]. Consider the *Etruscan from Marsiliana* symbol ▤. Revesz assumed that symbol has the vertical axis of symmetry in the letter, thus, we adopted this approach and related to this symbol the vertical axis of symmetry, denoted S_1 in Appendix and Supplementary Materials. We used the Schoenflies notation, for labeling the elements of symmetry of the symbols [15]. Unlike Revesz, we took into account all the symmetry elements (see Appendix and Supplementary Materials). Sometimes the decision about the presence or absence of an element of symmetry carries an inevitable element of subjectivity; for example, the letters ⸙ and ⸀ (*Etruscan from Marsiliana*, see Appendix and Supplementary Materials): for the first symbol we assume that there is horizontal symmetry, however, for the second we adopt that there is not.

The addressed alphabets are presented in Table 1.

Table 1. The alphabets of Canaanite group: Phoenician, Etruscan from Marsiliana, Archaic Etruscan, Neo-Etruscan, Archaic Western Greek, Euclidian Greek, Hebrew.

Phoenician	⼪	⼳	⼱	⼪	⼳	Y	I	⽇	⊗	⼱	⼪	∠	⼲	⼲	‡	O	⼲	⼲	Φ	⼪	W	X				
Etruscan Marsiliana	A	⼳	)	(I	⼳	⼳	I	⽇	⊗	I	⼪	⼴	⼲	⼲	⊞	O	⼲	M	Φ	⼪	S	T	Y	X	Φ	Y
Archaic Etruscan	A		)		⼳	⼳	I	⽇	⊗	I	⼪	⼴	⼲	⼲			⼲	M	Φ	⼪	S	T	Y	X	Φ	Y	8
Neo Etruscan	A		)		⼳	⼳	I	⽇	⊙	I		⼴	M	M			⼲	M		⼪	S	⼲	V		⊕	V
Western Greek	A	B	Γ	Δ	⼳	F	I	⽇	⊗	I	K	∧	⼲	⼲		O	Γ	∧	Φ	P	Σ	T	Y	X	Φ	Y
Euclidean Greek	A	B	Γ	Δ	E	Z	H	Θ	I	K	Λ	M	N	Ξ	O	Π	P	Σ	T	Υ	Φ	X	Ψ	Ω	A	B	Γ
Proto-Hebrew	⼪	⼳	⼲	⼪	⼳	⼲	I	⽇	⊗	⼱	⼪	L	⼲	⼲	‡	O	⼲	⼲	Φ	⼪	W	X				
Hebrew	א	ב	ג	ד	ה	ו	ז	ח	ט	י	כ	ל	מ	נ	ס	ע	פ	צ	ק	ר	ש	ת				
Archaic Latin	A	B	(	D	E	F	I	⽇	I	k	L	∧∧	∧I	O	Γ	Q	P	S	1́	VV	X					
Classical Latin	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	V	X	Y	Z			
English	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

For each of the studied alphabets, a table was compiled (see Appendix and Supplementary Materials), where for each character a row was filled in indicating the presence (labeled "1") or absence (labeled "0") of a given symmetry element, classified within the Schoenflies scheme. The exact shape of letters was taken as it is supplied in refs. [14,16,17,19,20,22].

3. Results

3.1. Quantitative characterization of the symmetry of alphabets; Shannon measures of symmetry and diversity of alphabets.

The genetic tree of the alphabets rooted in the Phoenician one is supplied in Figure 1 [18]. The supplied genetic tree is disputable, and we will address, at least partially, the problems related to its structure.

Figure 1. The Genetic tree of alphabets rooted in the Phoenician Script, as summarized in [16,19,20].

Symmetries of the symbols were analyzed and characterized with the Schoenflies notation as it is shown in Table A1 and Supplementary Materials. Let us illustrate the entire procedure with the Phoenician alphabet (for the symbols inscribed into a square) taken as an example. The symbols with only identity transformation (C_1) symmetry are:

⼳, ⼱, ⼳, ⼳, ∠, ⼲, ⼲, ⼲, ⼲, ⼪, Δ, ⼱ .

The symbols with identity transformation (C_1) and horizontal mirror axis (S_1) only is ⼪. The symbols with identity transformation (C_1) and vertical mirror axis (S_2) only are: Y, Φ, W. The symbols with identity transformation (C_1), horizontal and vertical mirror axes (S_1, S_2) and rotation on

180° (C_2) only are: ⌶, 🗄, ♯, ⬭. The symbols with identity transformation (C_1), horizontal, vertical and diagonal mirror axes (S_1, S_2, S_3, S_4) and 4-fold rotational symmetry (C_4, C_2, $C_4{}^3$) only are: ⊗, ✖.

Afterwards, two various Shannon symmetry measures were calculated for the addressed alphabets. The first is the Shannon/informational measure of symmetry of the alphabet $H_{\text{SYM}}(G)$ (abbreviated as IMS) defined in a Shannon-like form as follows:

$$H_{\text{SYM}}(G) = -\sum_{i=1}^{k} P_i(G_i) \ln P_i(G_i), \tag{1}$$

$$P_i(G_i) = \frac{m(G_i)}{N_G}, \tag{2}$$

where $P_i(G_i)$ is the probability of appearance of the symmetry operation G_i within the alphabet, and $N_G = \sum_{i=1}^{k} m(G_i)$ is the total number of the symmetry elements (operations) appearing in the alphabet and $m(G_i)$ is the number of the same symmetry elements/operations G_i, calculated for a given set of symbols/alphabet. The normalization condition given by Equation (3) takes place:

$$\sum_{i=1}^{k} P_i(G_i) = 1 \tag{3}$$

Table 2 summarizes $m(G_i)$ as established for the Phoenician Script; the total number of the elements of symmetry established for the Phoenician Script $N_g = 52$.

Table 2. Elements of symmetries inherent for the Phoenician alphabet and their frequencies and probabilities (symbols inscribed in a square).

Symmetry element, G_i	Elements of symmetry the Phoenician alphabet							
	C_1	S_1	S_2	S_3	S_4	C_4	C_2	$C_4{}^3$
$m(G_i)$	22	7	9	2	2	2	6	2
$P(G_i)$	0.423	0.135	0.173	0.038	0.038	0.038	0.115	0.038

Substitution of data, appearing in Table 2, and calculation with Eq. 1 yields: $H_{SYM}(Phoenician) = 1.688$. The aforementioned procedure was repeated for Western Greek, Euclidian Greek, Etruscan from Marsiliana, Archaic Etruscan, Neo-Etruscan, Proto-Hebrew, Hebrew, Archaic and Classic Latin and modern English scripts. The calculation was carried out for the symbols inscribed in a square (*s*-scripts) and also for the symbols inscribed into a rectangle (*r*-scripts). The symbol ⬭ was always considered with rectangle symmetry.

The second Shannon-like measure calculated for the addressed alphabets, known as Shannon diversity index (abbreviated SDI), we denote D_{sym}. For its calculation, we divide the total set of symbols/letters constituting the alphabet into subsets of symbols characterized by the same symmetry group (the same set of the symmetry operations). Shannon diversity index $D_{SYM}(\widetilde{G})$ is calculated as follows:

$$D_{SYM}(\widetilde{G}) = -\sum_{i=1}^{k} P_i(\widetilde{G}_i) \ln P_i(\widetilde{G}_i) \tag{4}$$

$$P_i(\widetilde{G}_i) = \frac{\widetilde{m}(\widetilde{G}_i)}{\widetilde{N}_{\widetilde{G}}}, \tag{5}$$

where $P_i(\widetilde{G}_i)$ is the probability of finding a subset of symbols with the same set of symmetry operations/symmetry group $\widetilde{G}_i$, $\widetilde{m}(\widetilde{G}_i)$ is the number of letters possessing the same symmetry group $\widetilde{G}_i$, $\widetilde{N}_{\widetilde{G}}$ is total number of subsets, which coincides with the number of letters in a given alphabet. Again, the normalization condition given by Equation (6) takes place:

$$\sum_{i=1}^{k} P_i(\widetilde{G}_i) = 1 \tag{6}$$

To calculate the Shannon diversity index, it is necessary to consider all subsets of symbols appearing in the alphabet characterized by the same set of symmetry elements. Table 3 shows these subsets with the probability of their appearance in the Phoenician alphabet (letters are inscribed in a square, $\widetilde{N}_{\widetilde{G}} = 22$). Substitution of these data into Eq. 4 yields: $D_{SYM}(\widetilde{G}) = 1.271$.

Table 3. Dividing the total set of symbols/letters constituting the Phoenician alphabet into the subsets of symbols characterized by the same symmetry group.

Subsets of symmetry elements, $\widetilde{G}_i$	C_1	$C_1 S_1$	$C_1 S_2$	$C_1 S_1 S_2 C_2$	$C_1 S_1 S_2 S_3 S_4 C_4 C_2 C_4{}^3$
Number of the letters. $\widetilde{m}(\widetilde{G}_i)$	12	1	3	4	2
Probability, $P_i(\widetilde{G}_i)$	0.545	0.045	0.136	0.182	0.091

Let us address graphs depicted in Figure 2 representing H_{SYM} calculated for the studied scripts and the plotted *vs.* the date at which the scripts were registered first [16,18–20,22]. For example, the first Etruscan text is ascribed to ca. 700 BCE [16]. The archaic Latin is ascribed to *ca.* 750 BCE [19,20]. Classical Latin is formed approximately at I century BCE {19-20}. The graphs established for the scripts inscribed into the square (represented with grey circles and abbreviated *s*-scripts) and the graphs established for the letters/graphemes inscribed into rectangles (abbreviated *r*-scripts) are depicted. It is recognized from the graphs presented in Figure 2 that H_{SYM} calculated for the *s*-scripts, is decreased with time; whereas H_{SYM}, is only slightly time sensitive for *r*-scripts. Let us explain this result: the average uncertainty to find an element of symmetry within the symbols of the given alphabet (averaged over the entire alphabet) is decreased with time for *s*-scripts, and it is constant for the *r*-scripts. The interpretation of this conclusion needs some care; indeed, $H_{SYM}(G) = -\sum_{i=1}^{k} P_i(G_i)\ln P_i(G_i)$, is not a monotonic function of P_i , indeed $H_{SYM}(G) = 0$, when $P_i = 0$, and also $H_{SYM}(G) = 0$, when $P_i = 1$ [21]. A low value of $H_{SYM}(G)$ may evidence, the absence of symmetry in the letters of the alphabet; and this the case with the Hebrew alphabet, for which H_{SYM} calculated for both *s*- and *r*-Hebrew scripts is very low and it is out of the entire picture. H_{SYM} calculated for modern English letters and supplied for the comparison in Figure 2 is close to those established for the Phoenician-rooted scripts.
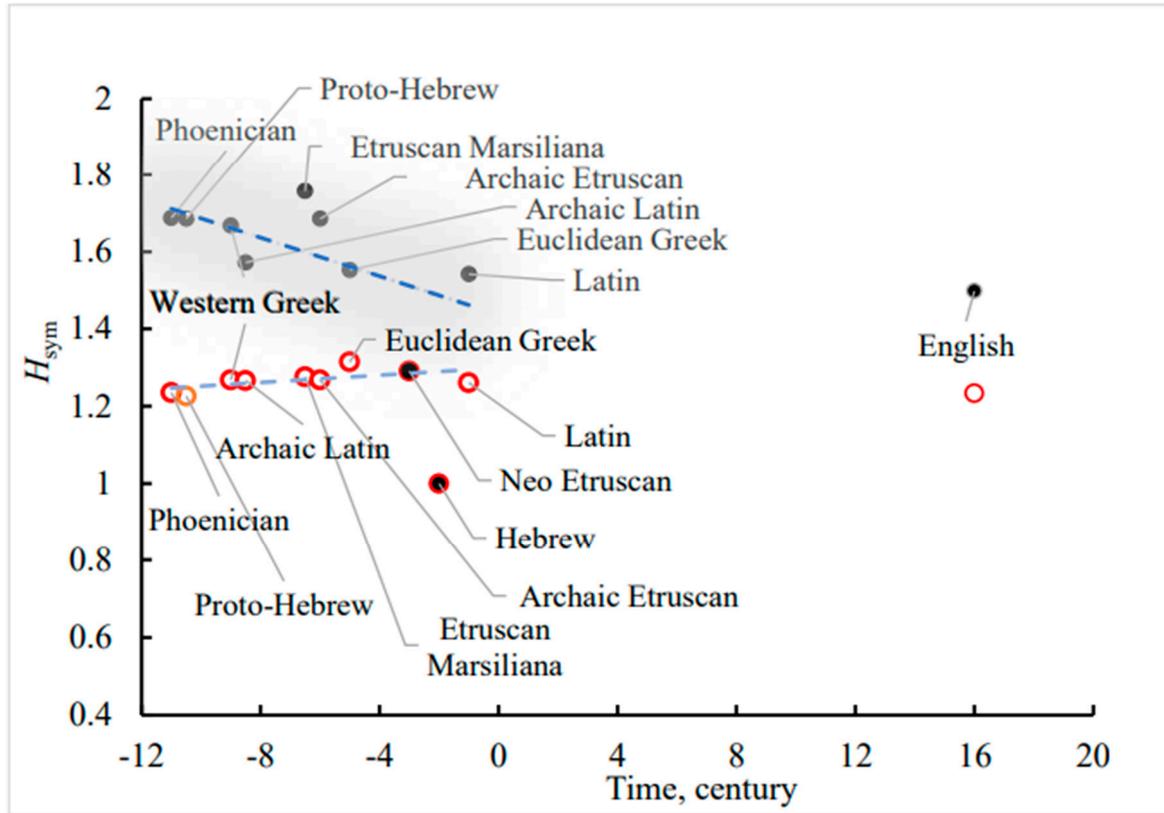
Figure 2. Time evolution of the informational measure of symmetry (IMS) calculated for various scripts: grey circles correspond to the letters inscribed into square (*s*-scripts); red circles correspond to letters inscribed into the rectangle (*r*-scripts).

Least squares regression straight lines emerging from the data plotted in Figure 2 are supplied by Eq. 7 and Eq. 8.

$$H_{sym}^{square}(t) = -0.0249 \times t + 1.437;\ R^2 = 0.3695\ , \tag{7}$$

$$H_{sym}^{rect}(t) = 0.0047 \times t + 1.2991;\ R^2 = 0.3593\ , \tag{8}$$

where R^2 is the squared correlation coefficient, which is calculated for all off the represented scripts without English and Hebrew/Ashurit which are obviously far from the regression trend line. The low values of the correlation coefficient evidence the fact that the straight lines are supplied for eye guidance only.

It is recognized from Eq. 7 and Eq. 8 that the modulus of the slope of the $H_{sym}^{rect}(t)$ regression line for *r*-scripts, is much lower than that established for *s*-scripts.

It is noteworthy that both H_{sym}^{square} and $H_{sym}^{rect}(t)$ are restricted within a very narrow range of values, namely: $1.291 < H_{sym}^{square} < 1.757$ and $1.226 < H_{sym}^{rect}(t) < 1.314$. The only exception is Hebrew $H_{sym}^{square} \cong H_{sym}^{rect}(t) \cong 1$. This observation will be discussed later.

Now we address the Shannon diversity index (SDI), denoted $D_{SYM}(\tilde{G})$ calculated for the studied alphabets with Eq. 4, and illustrated with Figure 3. Somewhat surprisingly, SDI is increased with time in a monotonic way for both of *s*- and *r*-scripts. This means that the diversity of symmetry groups inherent for alphabets emerging from the Phoenician one grows with time. And again, the Hebrew script, demonstrating markedly low value of D_{SYM} is an exception (see Figure 3). D_{SYM} calculated for modern English alphabet and supplied for the comparison in Figure 3 is close to those

established for ancient Phoenician-rooted scripts. Figure 3. Shannon diversity index (SDI), denoted $D_{SYM}(\tilde{G})$ calculated for the studied alphabets, rooted in the Phoenician one
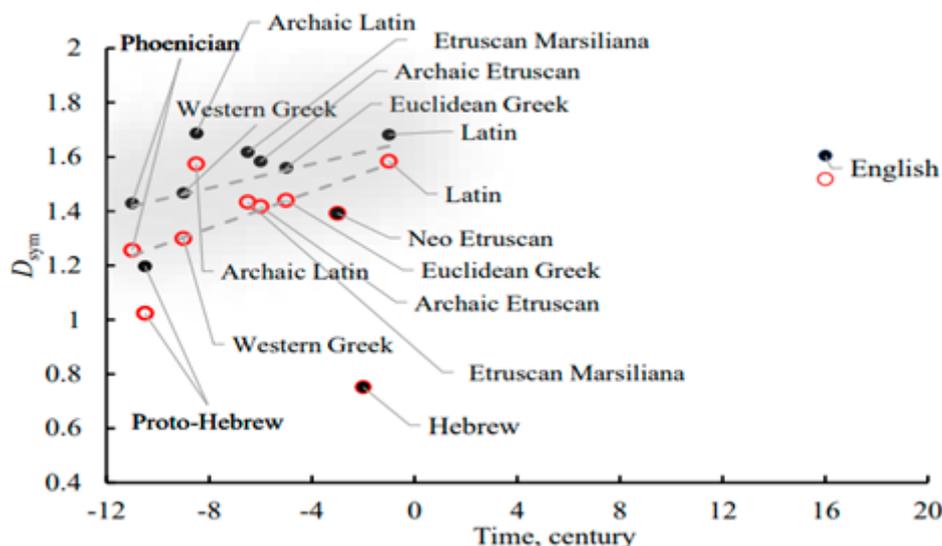

Figure 3. Shannon diversity index (SDI), denoted $D_{SYM}(\tilde{G})$ calculated for the studied alphabets, rooted in the Phoenician one

Least squares regression straight lines ($D_{sym}^{square}(t) = \alpha t + \beta$), emerging from the data plotted in Figure 3 are supplied by Eq. 9 and Eq. 10:

$$D_{sym}^{square}(t) = 0.0221 \times t + 1.6615; \; R^2 = 0.2233, \tag{9}$$

$$D_{sym}^{rect}(t) = 0.0337 \times t + 1.6066; \; R^2 = 0.4384, \tag{10}$$

The squared correlation coefficient R^2 is calculated for all off the represented scripts without English and Hebrew/Ashurit which are far from the trend line. It is recognized from Eq. 7 and Eq. 8 that the modulus of the slope of the $H_{sym}^{rect}(t)$ regression line for *r*-scripts, is much lower than that established for *s*-scripts. Again, low values of the correlation coefficient point to the fact that the straight lines are supplied for eye guidance only.

Let us take a more close glance on Eq. 9 and Eq. 10; as we already mentioned both of the dependencies $D_{sym}^{square}(t)$ and $D_{sym}^{rect}(t)$ grow with time; moreover, the slopes of the both of dependencies are of the same order of magnitude: $\alpha_{sym}^{square} \cong 0.021$; $\alpha_{sym}^{rect} \cong 0.037$. Let us calculate the points of intersection of the regression lines with the time axis: $D_{sym}^{square}(t) = 0.0221 \times \tau_{sym}^{*sq} + 1.6615 = 0$ and $D_{sym}^{rect}(t) = 0.0337 \times \tau_{sym}^{*rect} + 1.6066 = 0$. We calculate: $\tau_{sym}^{*sq} = -75.2 \; century$; $\tau_{sym}^{*rect} = -47.7 \; century$. The value $\tau_{sym}^{*rect} = -47.7 \; century$ catches the eye, due to the fact that it falls within the Vinča culture period, or Vinča-Turdaş culture, which is a Neolithic archaeological culture of Southeast Europe, dated to the period 5400–4500 BC [24–26]. The Vinča culture is a later Neolithic/early Chalcolithic phenomenon which lasted for 700 years in the largest part of the northern and central Balkans, spreading across an area which includes present-day Serbia, the Romanian Banat, parts of Romanian Oltenia, western Bulgaria, northern Macedonia and eastern parts of Slavonia and Bosnia [24–26].

Just to this period famous Vinča symbols (also called the Vinča script) are attributed [27]. The Vinča symbols, shown in Figure 4, are a set of untranslated symbols found on Neolithic era artifacts from the Vinča culture [27]. Whether this is one of the earliest writing systems or simply symbols of some sort is disputed [27]. Scholars have tried to answer two main questions about the nature of the signs: first, do they form a system, and (if so), could such a system be interpreted as an original prehistoric script? The scientists demonstrated that the signs and sign groups of the Vinča script are

uniform, just as in organized writing [27]. And it is reasonable to suggest, such a complex notation system could have been a form of written communication throughout the Vinča society. We plan to study the symmetry of Vinča symbols in our future investigations.

Figure 4. Vinča symbols appearing one of the Tărtăria tablets unearthed near Tărtăria, Romania, and dated to *ca*. 5.300 BCE.

Thus, if we speculate that the diversity of alphabets, constituting scripts, quantified by D_{sym}^{rect} and calculated with Eq. 4, evolved in time in a continuous wave, as shown in Figure 3, the regression line is expected to cross the axis of time in a point, to which the origin of scripts is related. And to the best of our knowledge in archeology this point of time coincides with the existence of the Vinča culture [24–27]. We are well aware that at this stage of investigation this is a bold hypothesis, which calls for the further investigations. The low value of the correlation coefficient of the linear regression appearing in Eq. 4, oblige con consider the aforementioned reasoning *cum grano salis*.

One more observation is noteworthy: the values of SDI are restricted in a narrow range of values for both *s*- and *r*-scripts $1.024 < SDI < 1.686$, with the only exception of the Hebrew script, namely $SDI \cong 0.752$.

3.2. Symmetry factor; its definition and calculation for alphabets

Now we introduce one more notion, enabling quantification of symmetry of symbols constituting the scripts. We adopt the plausible hypothesis that an amount of graphical information, necessary for storing the symbol is proportional to the area of rectangle in which the symbol may be inscribed. Mirror axes of symmetry will separate the rectangle into sub-areas, as shown in Figure 5. Consider symbol ⵦ - qoph of the Phoenician script. This symbol has the vertical mirror axis of symmetry denoted S_2 as depicted in Figure 5A. Thus, the entire symbol may be obtained by projection of half-a-symbol relatively the axis S_2 as shown in Figure 5B. If we have the full list of instructions describing building/drawing of half-a-symbol, the symmetrical projection will enable inscribing of the entire symbol. Thus, symmetry enables parsimony of information, necessary for drawing/inscribing of the symbols. Now consider the Phoenician letter ⊗ - teth , depicted in Figure 5C. This symbol has four mirror symmetry axes, namely (S_1, S_2, S_3, S_4), shown in Figure 5C. These axes separate the symbol into eight sub-segments, depicted in Figure 5C. Following the aforementioned reasoning, axes (S_1, S_2, S_3, S_4) provide the eight-fold parsimony of graphical information necessary for drawing/inscribing the symbol.

This eightfold parsimony of information may be also demonstrated with the Cayley table of symmetry of symbols [28,29]. It should be mentioned that for the ⊗ - theth symbol, we recognize four additional elements of symmetry and they are rotations about the geometrical center of the symbol to the angles $\varphi_i (i = 1 \dots 4) = \left(0; \frac{\pi}{2}; \pi; \frac{3\pi}{2}\right)$.
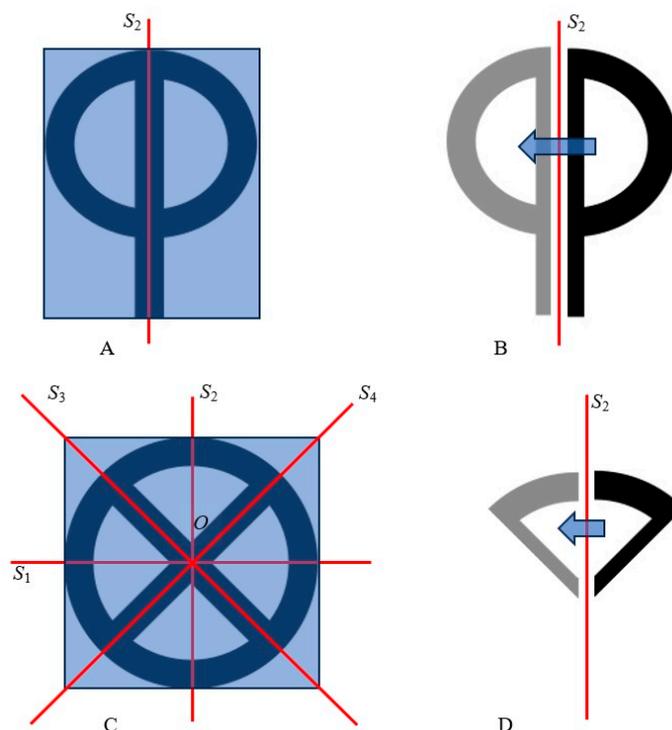
Figure 5. Analysis of symmetry of the symbols of Phoenician alphabet. **A**. Symbol.

⚵ - qoph   is depicted; mirror axis of symmetry S_2 is shown; **B**. The entire symbol may be obtained by projection of half-a-symbol relatively the axis S_2; thus, the two-fold parsimony of   information is provided. **C**. Symbol ⊗ - teth is shown. The symbol has four mirror   symmetry   axes,   namely (S_1, S_2, S_3, S_4) shown in the inset. D. The entire symbol may be   restored   by   the projection of the sub-segment, depicted in the inset; thus the eight-fold parsimony of information us provided.

Thus, the group of the symmetry of the symbol contains eight elements, namely four mirror axes and four distinguishable rotations [28,29]. Assume, that the letters are created with the software. Eight elements of symmetry provide an eightfold decrease in graphical information, necessary for drawing/inscribing the symbol. The same reasoning works for the Phoenician symbol ⚵ - qoph depicted in **Figure 5A**. The total symmetry group of this symbol contains the mirror axis S_2 and the identity element which is the rotation to $\varphi = 0$; thus, the total number of the symmetry operations is two. Hence, the symmetry provides the twofold parsimony of the graphical information necessary for drawing the symbol. It should be emphasized that the aforementioned reasoning does not depend on the specific type of drawing of the symbol. Now let us quantify the aforementioned parsimony. We denote $m_i(G)$ the total number of elements of symmetry related to i-th letter of the given alphabet, known in the group theory at the order of the group G [30]. Now we introduce the symmetry factor of the alphabet denoted μ and defined with Eq. 11:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{m_i(G)} \quad , \tag{11}$$

where n is the number of symbols in the alphabet. The symmetry factor μ quantifies the averaged level of symmetrization of the alphabet on one hand, and the possible parsimony of graphical information necessary for the drawing of the entire set of letters, constituting the alphabet. **Figure 6** depicts the dependence of the symmetry factor μ claculated for various alphabets of the Phoenician group.
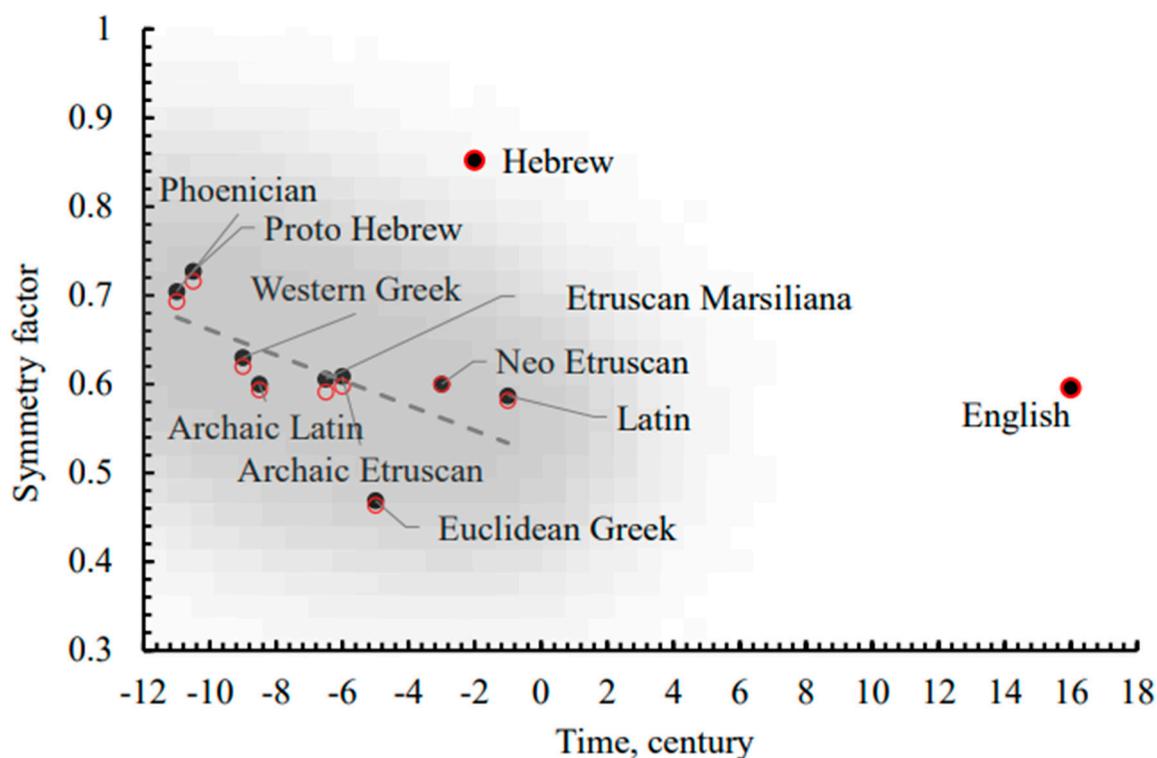
Figure 6. Symmetry factor μ calculated for the Phoenician branch alphabets; solid circles correspond to the letters inscribed into square (*s*-scripts); red circles correspond to the into the rectangle (*r*-scripts).

The regression line describing the time evolution of the symmetry factor $\mu(t)$ is given by Eq. 12:

$$\mu(t) = -0.0142 \times t + 0.5195; \; R^2 = 0.4188 \tag{12}$$

The regression line crosses the time axis at the point $\tau^* = 36.6\; century$. Let us take a close look on the plot, presented in Figure 6. We come to the following conclusions: i) points representing rectangular and square scripts are located very close one to another for all of the studied scripts, emerging from the Phoenician alphabet; ii) symmetry factor μ is decreased with time. This means, that the averaged level of symmetrization of the studied alphabet is increased with time; and the parsimony of graphical information necessary for writing is increased with time. And, again the high value of the symmetry factor established for Hebrew presents the obvious exception.

4. Discussion

Language is a system that we rely on for both interpersonal and intrapersonal communication [31]. It is also one of the core elements of any culture or human civilization [31]. Language is a hallmark that distinguishes human beings from other species [31]. An African tradition has a keen insight into this aspect of language when people in a certain region of Africa call a newborn child *a kintu*, a "thing," until the child acquires a language. Once the child acquires the mother tongue, he/she can become *a muntu*, a "person" [31]. The fundamental question is: does the language we speak and write shape the way we think about the world? The problem is extremely perplexed; however, it was hypothesized that the alphabet promotes linear thinking and hierarchical reasoning, such as Aristotelian Syllogism [32].

In our paper we addressed the time evolution of alphabets, rooted in the Phoenician one. More rigorously speaking, we tried to quantify the time evolution of symmetry of the symbols, constituting the scripts emerging from the Phoenician alphabet. Of course, the study of symmetry is not the

unique way of investigation of the evolution of alphabets, but it is one of the possible pathways.   The word alphabet was originally derived from the Semitic "alphabet" whose first and second letters were "aleph" (meaning *ox)* and "bayit" (meaning *house*), respectively. Based on these two graphemes, the first two letters of the Greek alphabet, "alpha" and "beta", were created, which in turn became the word alphabet. The Phoenician alphabet served the source for two Semitic alphabets: the early Hebrew alphabet and the Aramaic alphabet. These two alphabets used the Phoenician alphabet at first, but the people developed their own national characters, beginning in 850 B.C. for Hebrew and 750 B.C. for Aramaic, and kept 22 letters of the Phoenician alphabet [31,32].

Our investigation is focused on the quantitative study of symmetry of letters constituting alphabets emerging from the Phoenician one. Why the symmetry of letters is important? It is important for two reasons: i) symmetry impacts the perception of letters [33–36]; ii) symmetry is important from the point of view of effort necessary for creation/drawing of letters. It was demonstrated, that there exists the relation between our perception of symbols and their symmetry. Findings indicate that while visual complexity is an obstacle to be overcome for early learners, the expert viewer is able to exploit complexity for improved performance [33]. Visual complexity is not synonymous to symmetry, but it is closely related to it [33–36].

We addressed in our investigation another aspect in the role of symmetry in emergence of alphabets. It seems intuitively clear, that symmetrical letters are easier to draw. How the effort necessary for their drawing may be quantified? We suggest such a measure labeled in our paper the "symmetry factor", which is defined with Eq. 11. We demonstrate that symmetrical letters enable parsimony of information necessary for drawing of graphemes/letters, which are the smallest functional units of a script.

We also investigated the time evolution of quantitative measures of symmetry, calculated for the letters, constituting the alphabets related to Phoenician group of scripts. The symmetry of the symbols of Phoenician Scripts was investigated by Revesz [14]. It was shown that many scripts in the Phoenician Script family contain a high percentage of signs that have mirror symmetry [14]. Moreover, the scripts within the Phoenician script family show a tendency of increased percentage of mirror-symmetric signs over time [14]. For example, while the Phoenician Alphabet contains 40.9 percent mirror-symmetric signs, one of its descendants, the Euclidean Greek Alphabet, contains 59.3 percent mirror-symmetric signs [14]. Revesz identified the boustrophedonic way of writing and religious writings with a deliberate mirroring as an afterlife symbolism as possible causes of the increased use of mirror symmetric signs [14]. In our study, we considered all kinds of symmetry appearing in the Phoenician scripts.

We demonstrate that the Shannon diversity index of symmetry is increased with time in a monotonic way for the studied scripts. This means that the diversity of symmetry groups inherent for alphabets emerging from the Phoenician one grows with time.

It should be mentioned that the Modern Hebrew script also called also Ashurit or square script presents an obvious exception from the entire set of Phoenician scripts: quantitative measures of symmetry calculated for this scripts are very different from those established for other Phoenician rooted scripts. It remains unclear, whether the Ashurit script is rooted in Phoenician script. [37] Our study supplied the arguments for the point of view, according to which the modern Hebrew/Ashurit script does not emerge from the Phoenician one [37}.

5. Conclusions

Study of the time evolution of symmetry of letters constituting alphabets emerging from the Phoenician script is reported. A diversity of quantitative measures of symmetry of graphemes appearing in Phoenician, Etruscan from Marsiliana, Archaic Etruscan, Neo-Etruscan, Archaic Western Greek, Euclidian Greek, Archaic and Classical Latin, Proto-Hebrew and Hebrew/Ashurit scripts, constituting the Phoenician script family, were calculated. The Shannon-like measures of symmetry were calculated. The Shannon measure of symmetry, denoted H_{SYM} was calculated for the scripts inscribed into the square (abbreviated *s*-scripts) and also for the scripts inscribed into rectangles (abbreviated *r*-scripts) [38]. H_{SYM}, calculated for the *s*-scripts, is decreased with time;

whereas H_{SYM} is only slightly time sensitive for r-scripts. We also calculated the Shannon diversity index (SDI), denoted $D_{SYM}(\tilde{G})$ for the s- and r-scripts. SDI is increased with time in a monotonic way for both of s- and r-scripts. This means that the diversity of symmetry groups inherent for alphabets emerging from the Phoenician one grows with time. And this is a very important conclusion. We also introduced the symmetry factor of the alphabet denoted μ and defined as follows: $\mu = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i(G)}$, where $m_i(G)$ the total number of elements of symmetry related to i-th letter of the given alphabet and n is the number of graphemes in the alphabet. The symmetry factor μ quantifies the averaged level of symmetrization of the alphabet on one hand. This factor also quantifies the possible parsimony of graphical information necessary for the drawing of the entire set of graphemes, constituting the alphabet on the other hand. The symmetry factor is decreased with time for the alphabets rooted in the Phoenician one. This means, that the averaged level of symmetrization of the studied alphabet is increased with time; thus, resulting in simplification of writing. In other words, the parsimony of graphical information necessary for writing graphemes is increased with time. We also concluded that the values of the symmetry factor calculated for rectangular and square scripts are close one to another for all of the studied scripts, emerging from the Phoenician alphabet, with the pronounced exception of the Hebrew/Ashurit script. Thus, our study supplies the arguments for the point of view, according to which the modern Hebrew/Ashurit script does not emerge from the Phoenician one [37]. Why the symmetry of graphemes is important?

It is important for a number of reasons: i) symmetry impacts the perception of letters; ii) symmetry is important from the point of view of effort necessary for creation/drawing of letters, iii) symmetry to a much extent quantifies the "orderliness" of the pattern built of the graphemes [39,40]. Orderliness of the pattern built of the graphemes/text has a fine structure and it could not be quantified by a single numerical value; the symmetry factor μ is one of the parameters describing the order in a given pattern [39,40]. It should be emphasized that the Shannon Symmetry Measure and the Shannon Diversity Factor are probabilistic measures of symmetry in their nature; whereas, the symmetry factor μ is not.

We conclude that the study of time evolution of symmetry of scripts supplies valuable information about their functioning as graphical systems and origin.

Appendix

Table A1. Symmetry elements for each letter of the Phoenician alphabet classified within the Schoenflies scheme. The presence of the symmetry element is labeled as "1", absence – as "0". "1/0" means that symmetry element is present in square configuration, but not in rectangular one.

Phoenician alphabet							
Script	C_1	S_1	S_2	S_3	S_4	C_4	C_2	C_4^3
	1	1	0	0	0	0	0	0
	1	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0

Symbol								
ⱦ	1	0	0	0	0	0	0	0
ꓯ	1	0	0	0	0	0	0	0
ꓶ	1	0	1	0	0	0	0	0
Δ	1	1	1	0	0	0	1	0
ꓱ	1	1	1	0	0	0	1	0
Y	1	1	1	1/0	1/0	1/0	1	1/0
I	1	0	0	0	0	0	0	0
ꓧ	1	0	0	0	0	0	0	0
⊗	1	0	0	0	0	0	0	0
ꓺ	1	0	0	0	0	0	0	0
ⱶ	1	0	0	0	0	0	0	0
L	1	1	1	0	0	0	1	0
ꟽ	1	1	1	0	0	0	1	0
ꓭ	1	0	0	0	0	0	0	0
ꓩ	1	0	0	0	0	0	0	0
O	1	0	1	0	0	0	0	0
ꓷ	1	0	0	0	0	0	0	0
�	1	0	1	0	0	0	0	0
Φ								
ꓭ	1	1	1	1/0	1/0	1/0	1	1/0
W								
X								

References

1. Sampson, G. Writing systems, in: Methods for recording language. Chapter 4., pp. 47-62, in The Routledge Handbook of Linguistics, ed. by Keith Allan, 2016, Routledge, 711 Third Avenue, New York, NY 10017, USA.

2. Keith, A. What is linguistics? in: Methods for recording language. Chapter 4., pp. 1-16, in The Routledge Handbook of Linguistics, ed. by Keith Allan, 2016, Routledge, 711 Third Avenue, New York, NY 10017, USA.

3. Gardiner, A. H, The egyptian origin of the semitic alphabet. J Egypt Archeol. **1916**, *3*(1), 1–16.

4. Pennacchietti, F. A. An alternative hypothesis on the origin of the Greek alphabet, Kervan, International Journal of African and Asiatic Studies, 2023, 27(1), 109-121.

5. Powell, B. B. Homer and the Origin of the Greek Alphabet, Ch. 1, pp. 5-66, 1991, Cambridge, Cambridge University Press.

6. Hurford, J. R. Evolutionary Linguistics: How Languages and Language Got to Be the Way They Are: in: Methods for recording language. Chapter 2., pp. 17-32, in The Routledge Handbook of Linguistics, ed. by Keith Allan, 2016, Routledge, 711 Third Avenue, New York, NY 10017, USA.

7. Weyl, H. Symmetry; Princeton University Press: Princeton, NJ, USA, 1989.

8. Garrod, S.; Fay, N.; Lee, J.; Oberlander, J.; Macleod, T. Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, **2007**, *31*(6), 961–987.

9. Kelly, P. The invention, transmission and evolution of writing: insights from the new scripts of West Africa. In Paths into script formation in the ancient Mediterranean. Silvia Ferrara and Miguel Valério, eds. pp. 189–209. 2018, Rome: Studi Micenei ed Egeo Anatolici.

10. Kelly, P.; Winters, J.; Miton, H.; Morin, O. The predictable evolution of letter shapes: An emergent script of West Africa recapitulates historical change in writing systems, *Current Anthropology*, **2021**, *62* (6), 669-691.

11. Gibson, E.; Futrell, R.; Piantadosi, S. T.; Dautriche, I.; Mahowald, K., Bergen, L.; Levy, R. How efficiency shapes human language. *Trends in Cognitive Sciences*, **2019**, *23*, 389–407.

12. Han, S. J.; Kelly, P.; Winters, J.; Kemp, C. Simplification Is Not Dominant in the Evolution of Chinese Characters. Open Mind: *Discoveries in Cognitive Science*, **2022**, 6, 264–279.

13. Shannon, C.E.; Weaver, W. The Mathematical Theory of Communication; The University of Illinois Press: Chicago, IL, USA, 1949; Volume 97, pp. 29–51.

14. Revesz, P. R. The development and role of symmetry in ancient scripts, in: Viana, V., Nagy, D., Xavier, J., Neiva, A., Ginoulhiac, M., Mateus, L. & Varela, P. (Eds.). (**2022**). *Symmetry: Art and Science | 12th SISSymmetry Congress [Special Issue]*. Symmetry: Art and Science. International Society for the Interdisciplinary Study of Symmetry. 308-315.

15. Altmann, S. L.; Herzig, P. Point-group theory tables, Oxford,   Oxford University Press, UK, 1994.
16. Bonfante, J.H., Bonfante, L., P. The Etruscan Language: An Introduction, Manchester University Press, UK, 1983.
17. Stützer, H. A., Die Etrusker und ihre Welt,   DuMont DuMont Reiseverlag , Koln, Ge, 1992.
18. Fischer, S. R. History of writing, p. 298, Reaktion books, London, 2005.
19. Федорова, Е. В., Введение в латинскую эпиграфику, Изд-во Московского университета, 1982
20. Ярцева, В. Н., Большой энциклопедический словарь Языкознание, Изд-во "Советская энциклопедия", Москва, 1989
21. Ben-Naim, A. Entropy, Shannon's Measure of Information and Boltzmann's H-Theorem. Entropy 2017, 19, 48
22. Skolnik, F.; Berenbaum M., Encyclopaedia Judaica, Volume 1, 2nd ed., Thomson Gale ; Macmillan Reference USA, [Farmington Hills, Mich.], Detroit, 2007.
23. Jost, L Entropy and diversity, *Oikos*. **2006**, *113* (2), 363–375.
24. Amicone, S.; Radivojević, M.; Quinn, P.S.; Berthold, C.; Rehren, T. Pyrotechnological connections? Re-investigating the link between pottery firing technology and the origins of metallurgy in the Vinča Culture, Serbia, *J. Arch. Sci*. **2020**, *118*, 1-19.
25. Radivojević, M; Archaeometallurgy of the Vinča culture: A case study of the site of Belovode in Eastern Serbia. *Historical Metallurgy*, **2014**, *47* (1), 13-32.
26. Radivojević, M.; Rehren, T.; Pernicka, E.; Šljivar, D.; Brauns, M.; Borić, D., On the origins of extractive metallurgy: New evidence from Europe, *J. Archaeological Sci*. **2010**, *37*, 2775–2787.
27. Starović, A. If the Vinča script once really existed who could have written or read it? *Documenta Praehistorica* **2005**, *32*, 253–260.
28. Rosen, J. Symmetry in Science: An Introduction to the General Theory; Springer: Berlin, Germany, 1995.
29. Chatterjee, S.K. Crystallography and the World of Symmetry; Springer: Berlin, Germany, 2008.
30. Arfken G. B.; Weber H. J. Mathematical Methods for Physicists, 5ed, A Harcourt Science and Technology Company, San Diego, USA, 2001 .
31. Pae, H. K. Script Effects as the Hidden Drive of the Mind, Cognition, and Culture, Literacy Studies. Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education, v, 21. Springer, Cham, Switzerland, 2020.
32. Logan, R. K. The alphabet effect: a media ecology understanding of the making of Western civilization. Cresskill, NJ: Hampton Press, 2004.
33. Wiley, R. W., Wilson, C., & Rapp, B. The effects of alphabet and expertise on letter perception. *J. Exp. Psychology: Human Perception & Performance,* **2016**, *42* (8), 1186–1203.
34. Somov, G. U. Interrelations of codes in human semiotic systems, *Semiotica* **2016,** *213*, 557–599.
35. Yodogawa, E. Symmetropy, an entropy-like measure of visual symmetry. *Percept. Psychophys*. **1982**, *32*, 230–240.
36. Dry, M.J. Using relational structure to detect symmetry: A Voronoi tessellation based model of symmetry perception. *Acta Psychol*. **2008**, *128*, 75–90.
37. Hoffman, J. M. In the beginning. A Short History of Hebrew Language, New York University Press, New York, 2004.
38. Bormashenko, Ed.; Legchenkova, I.; Frenkel, M.; Shoval, Sh. Shannon (Information) Measures of Symmetry for 1D and 2D Shapes and Patterns, *Appl. Sci*. **2022**, *12*(3), 1127.
39. Bormashenko, Ed. Entropy, Information, and Symmetry: Ordered is Symmetrical, *Entropy* **2020**, *22*(1), 11.
40. Bormashenko, Ed. Entropy, Information, and Symmetry; Ordered Is Symmetrical, II: System of Spins in the Magnetic Field, *Entropy* **2020**, *22*(2), 235.