



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

Artem Gilevich¹, Mark Frenkel¹, Shraga Shoval² and Edward Bormashenko^{1,*}

- ¹ Department of Chemical Engineering, Ariel University, P.O. Box 3, Ariel 407000, Israel; artgilevich@gmail.com (A.G.); markfr@ariel.ac.il (M.F.)
- ² Department of Industrial Engineering and Management, Faculty of Engineering, Ariel University, P.O. Box 3, Ariel 407000, Israel; shraga@ariel.ac.il
- * Correspondence: edward@ariel.ac.il

Abstract: We investigated the time evolution of the symmetry of symbols constituting alphabets rooted in the Phoenician script. 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 established for the Hebrew/Ashurit and English scripts. The Shannon-like measures of symmetry were computed. The Shannon diversity index was calculated. Our findings indicate that the Shannon diversity index increased with time in a monotonic way for the studied scripts. The diversity of symmetry groups inherent for addressed alphabets 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. We found that the symmetry factor is decreased with time for the alphabets rooted in the Phoenician script. This means that the average level of symmetrization of the studied alphabet increases with time. The parsimony of graphical information necessary for writing graphemes is consequently increased with time. The values of the symmetry factor calculated for the addressed scripts are close to one 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 did 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 approximately 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 and sounds [2]. In our manuscript, we propose the physical/mathematical 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 the symbols, constituting the scripts, rooted in the Phoenician alphabet. For the 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.



Citation: Gilevich, A.; Frenkel, M.; Shoval, S.; Bormashenko, E. Time Evolution of the Symmetry of Alphabet Symbols and Its Quantification: Study in the Archeology of Symmetry. *Symmetry* 2024, *16*, 465. https://doi.org/ 10.3390/sym16040465

Academic Editors: Michel Planat, Sergei D. Odintsov and Alexei Kanel-Belov

Received: 13 December 2023 Revised: 19 March 2024 Accepted: 1 April 2024 Published: 11 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). As alphabets cannot develop without the evolution of a language, our research is partially related to evolutionary linguistics. "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; hence, 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].

Over the past decade, researchers have shown an increased interest in the time evolution of scripts. 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 the 19th 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 analysis of Chinese characters spanning more than 3000 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 may the "simplicity" or "information compression" 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 the alphabets emerging from the Phoenician alphabet [13]. We also demonstrate that the Shannon measure of symmetry and the Shannon diversity index are adequate mathematical tools, enabling a quantitative description of the symbols [13]. We also introduce the symmetry factor of the alphabet, which quantifies the average symmetry of its symbols. Our research further develops the approach suggested by Revesz, who highlighted the importance of symmetry for the analysis of the time evolution of alphabets [14,15]. Our investigation was motivated by the search for the adequate, quantitative, mathematical concept, describing the time evolution of alphabets. We demonstrate that the symmetry of graphemes may serve as the Ariadne thread, which unify and quantify such an evolution.

2. Methods

Scripts Addressed in the Investigation

In this study, scripts constituting the Phoenician script family were analyzed. Phoenician, Etruscan from Marsiliana, Archaic Etruscan, Neo-Etruscan [16], Archaic Western Greek, Euclidian Greek, Archaic, Classical Latin [17,18], and Proto-Hebrew scripts were addressed. The Hebrew/Ashurit script, whose relation to the Phoenician group is disputable, was also analyzed [19]. Scripts and shapes of the symbols are summarized in Appendix A and Supplementary Materials. Analysis of the symmetry of the Phoenician script family exploited the procedure, introduced originally by Revesz [14]. This procedure considered the vertical symmetry of symbols [14]. Consider the *Etruscan from Marsiliana* symbol **E**. Revesz assumed that this 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 A 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 A and Supplementary Materials). Sometimes, the decision about the presence or absence of an element of symmetry carries an inevitable element of subjectivity; for example, consider the letters **4** and **1** (*Etruscan from Marsiliana*, see Appendix A and Supplementary Materials). For the first symbol, we determine that there is horizontal symmetry; however, there is no horizontal symmetry for the second symbol. The addressed alphabets investigated in this study are presented in Table 1.

Phoenician	*	٩	1	٥	Ħ	Y	I	Β	\otimes	٦	ĸ	Z	*)	4	ŧ	0	2	٣	φ	٩	w	х					
Etruscan Marsiliana	A	8)	(Π	7	I	E	8	I	к	1	~1	٦	⊞	0	1	Μ	Ŷ	4	5	Т	٢	Х	φ	¥	
Archaic Etruscan	A)		Щ	7	I	B	8	١	к	1	~1	٦			1	Μ	Ŷ	٩	ζ	Т	٢	Х	φ	Ψ	8
Neo- Etruscan	A)		щ	7	Ι	Β	0	١		1	111	11			1	Μ		٥	ζ	ł	v		Φ	¥	
Western Greek	A	8	٢	۵	E	F	I	Β	8	١	K	\wedge	٣	۲		0	Г	Μ	Ŷ	٢	٤	Т	Y	Х	Φ	Ψ	
Euclidean Greek	А	В	Г	Δ	E	Z	Н	Θ	Ι	K	Λ	М	N	[1]	0	П	Р	Σ	Т	Υ	Φ	Х	Ψ	Ω			
Proto- Hebrew	≮	4		٩	E	Ч	Ι	Β	8	4	ሃ	L	Ч	Ч	ŧ	0	7	٣	φ	٩	W	×					
Hebrew	א	ב	ړ	т	ъ	I	r	Π	υ	ı	С	ל	מ	נ	0	ע	פ	Я	ק	٢	ש	л					
Archaic Latin	R	В	(D	E	F	I	В	I	k	L	W	N	0	Γ	Q	Р	\$	1	٧Y	Х						
Classical Latin	A	В	С	D	Е	F	G	Н	I	К	L	М	Ν	0	Ρ	Q	R	S	Т	V	Х	Y	Z				
English	А	В	С	D	Е	F	G	Н	Ι	J	К	L	Μ	Ν	0	Ρ	Q	R	S	Т	U	V	W	Х	Y	Ζ	

Table 1. The alphabets of the Canaanite group.

For each of the studied alphabets, a table was compiled (see the Appendix A 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 [15]. The exact shape of letters was taken as supplied in refs. [14,16–19].

3. Results

3.1. Quantitative Characterization of the Symmetry of Alphabets: Shannon Measures of Symmetry of the Alphabets

The genetic tree of the alphabets rooted in the Phoenician script is supplied in Figure 1 [20]. The supplied genetic tree is disputable, and we will address, at least partially, the problems related to its structure and origin.

The North Semitic alphabet is the earliest fully developed alphabetic writing system. It was used in Syria as early as the 11th century BC and is probably an ancestor, either directly or indirectly, of all subsequent alphabetic scripts, with the possible exception of those scripts classified as South Semitic (e.g., Ethiopic, Sabaean).

1300 BC	NORTH SEMIT	IC ALPHABET
11 BC	PHOENICAN ALPHABET	ARAMAIC ALPHABET
9 BC	GREEK ALPHABET	
7 BC	ETRUSCAN ALPHABET	HEBREW
1 AD	CLASSIC LATIN	
700AD	WESTERN EUROPIAN ALPHABET	

Figure 1. The genetic tree of alphabets rooted in the Phoenician script, as summarized in [16,17,19].

As previously mentioned, the symmetries of the symbols were analyzed and characterized with the Schoenflies notation, 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

The symbol with identity transformation (C_1) and horizontal mirror axis (S_1) only is \blacktriangleleft . The symbols with identity transformation (C_1) and vertical mirror axis (S_2) only are Υ , Φ , Ψ . The symbols with identity transformation (C_1), horizontal and vertical mirror axes (S_1 , S_2) and rotation on 180° (C_2) only are \mathbf{I} , \mathbf{E} , \mathbf{O} . The symbols with identity transformation (C_1), horizontal and vertical mirror axes (S_1 , S_2) and rotation on 180° (C_2) only are \mathbf{I} , \mathbf{E} , \mathbf{O} . 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 $\mathbf{\Theta}$, \times .

Following the introduced classification of symmetry elements, two different Shannon measures were calculated for the addressed alphabets. The first is the Shannon/informational measure of symmetry of the alphabet $H_{SYM}(G)$ (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 the appearance of the symmetry operation G_i within the alphabet, $N_G = \sum_{i=1}^k m(G_i)$ is the total number of symmetry elements (operations) appearing in the alphabet and $m(G_i)$ is the number of same symmetry elements/operations G_i , calculated for a given set of symbols/alphabet. The normalization condition given by Equation (3) takes place by:

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

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

Substitution of the data appearing in Table 2 and calculation with Equation (1) yields $H_{SYM}(Phoenician) = 1.688$. This 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 \circ was always classified with rectangle symmetry.

Elements of Symmetry the Phoenician Alphabet											
Symmetry Element, G	$G_i C_1$	<i>S</i> ₁	<i>S</i> ₂	<i>S</i> ₃	S_4	C_4	<i>C</i> 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			

Table 2. Elements of symmetries inherent for the Phoenician alphabet and their frequencies and probabilities (*s*-script).

3.2. Shannon Diversity Index of the Alphabets

The second Shannon-like measure calculated for the addressed alphabets is known as the Shannon diversity index (abbreviated SDI) [21], which we denote as D_{sym} . For this 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). The Shannon diversity index $D_{SYM}(\widetilde{G})$ is calculated as follows:

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

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

where $P_i(G_i)$ 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 and $\widetilde{N}_{\widetilde{G}}$ 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, $\tilde{N}_{\tilde{G}} = 22$). Substitution of these data into Equation (4) yields $D_{SYM}(\tilde{G}) = 1.271$.

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

Subsets of Symmetry Elements, \tilde{G}_i	<i>C</i> 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 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 the graphs depicted in Figure 2 representing H_{SYM} calculated for the studied scripts and the plotted vs. the date at which the scripts were first registered [16–20]. For example, the first Etruscan text is approximately ascribed to 700 BCE [16]. The archaic Latin is ascribed to ca. 750 BCE [17,18]. Classical Latin is formed approximately in approximately the 1st century BCE [17,18]. The graphs established for the scripts inscribed into the square (represented with gray 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) decreases 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 , and indeed $H_{SYM}(G) = 0$ when $P_i = 0$ and also $H_{SYM}(G) = 0$ when $P_i = 1$ [22]. A low value of $H_{SYM}(G)$ may illustrate 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 trend lines, shown in Figure 2. H_{SYM} calculated for modern English letters and supplied for the comparison in Figure 2 is not far from those established for the Phoenician-rooted scripts, and it is very close to H_{SYM} calculated for the Latin script.



Figure 2. Time evolution of the informational measure of symmetry (IMS) calculated for various scripts: gray circles correspond to the letters inscribed into a square (*s*-scripts); red circles correspond to letters inscribed into the rectangle (*r*-scripts). Two dash lines corresponds to square and rectangle scripts respectively.

Least squares linear regressions emerging from the data plotted in Figure 2 are supplied by Equations (7) and (8)

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

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

where R^2 is the squared correlation coefficient, which is calculated for all the represented scripts excluding 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 visual guidance only. The straight regression lines span the time period from 12 BC to 1 BC, when the time evolution of the addressed scripts stopped.

It is recognized from Equations (7) and (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 the *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 $D_{SYM}(G)$, calculated for the studied alphabets with Equation (4) and illustrated with Figure 3. Somewhat surprisingly, SDI is increased with time in a monotonic way for both of the *s*- and *r*-scripts. This means that the diversity of symmetry groups inherent for alphabets emerging from the Phoenician script grows with time. And again, the Hebrew script, demonstrating a markedly low value of D_{SYM} , is an exception (see Figure 3).



Figure 3. Shannon diversity index (SDI), denoted $D_{SYM}(G)$, calculated for the studied alphabets, rooted in the Phoenician one. Black circles correspond to the letters inscribed into a square (*s*-scripts); red circles correspond to letters inscribed into rectangles (*r*-scripts).

It should be mentioned that D_{SYM} calculated for the modern English alphabet and supplied for the comparison in Figure 3 is not far from those established for ancient Phoenician-rooted scripts, and it very close to that calculated for the Latin script. Thus, we conclude that the time evolution of the Phoenician-rooted scripts stopped with the appearance of the Latin script. Thus, the modern English script is excluded from the straight regression lines, describing the time evolution of the aforementioned scripts.

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

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

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

The squared correlation coefficient R^2 is calculated for all the represented scripts excluding English and Hebrew/Ashurit, which are far from the trend line. It is recognized from Equations (7) and (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 the *s*-scripts. Again, low values of the

correlation coefficient point to the fact that the linear regressions are supplied for visual guidance only.

3.3. Shannon Diversity Index of Alphabets and the Vinča Symbols

Let us take a close look at Equations (9) and (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 [23–25]. 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 [23–25].

The famous Vinča symbols (also called the Vinča script) are attributed to this period [26]. The Vinča symbols, shown in Figure 4, are a set of untranslated symbols found on Neolithic era artifacts from the Vinča culture [26]. Whether this is one of the earliest writing systems or simply symbols of some sort is disputed [26]. 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 [26]. And it is reasonable to suggest that 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*. 5300 BCE. The scale bar is 3 cm.

Thus, if we speculate that the diversity of alphabets, constituting scripts, quantified by D_{sym}^{rect} and calculated with Equation (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 the scripts is related. And to the best of our knowledge, archeology studies identify this point in time as coinciding with the Vinča culture [23–26]. We are well aware that at this stage of investigation that this is a bold hypothesis which calls for further investigations.

The low value of the correlation coefficient of the linear regression appearing in Equation (4) can lead to the aforementioned reasoning being considered 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, $SDI \cong 0.752$.

3.4. Symmetry Factor: Its Definition and Calculation for Alphabets

Now, we introduce one more notion, enabling quantification of the symmetry of the symbols constituting the scripts. We adopt the plausible hypothesis that the amount of graphical information necessary for storing/displaying the symbol is proportional to the area of a rectangle in which the symbol may be inscribed. Mirror axes of symmetry separate the rectangle into sub-areas, as shown in Figure 5. Consider symbol $\mathbf{\Phi}$ —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 over the axis S_2 , as shown in Figure 5B. If we have the full list of instructions describing the building/drawing of half a symbol, the symmetrical projection will enable the entire symbol to be inscribed. Thus, symmetry enables parsimony of information, necessary for drawing/inscribing of the symbols. Now, consider the Phoenician letter $\mathbf{\Phi}$ —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 reasoning, axes (S_1, S_2, S_3, S_4) provide the symbol.

This eightfold parsimony of information may also be demonstrated with the Cayley table of symmetry of symbols [27,28]. It should be mentioned that for the \otimes (theth) symbol, we recognize four additional elements of symmetry, which are rotations about the geometrical center of the symbol to the angles $\varphi_i(i=1...4) = (0; \frac{\pi}{2}; \pi; \frac{3\pi}{2})$. Thus, the group of the symmetry of the symbol contains eight elements, namely, four mirror axes and four distinguishable rotations [27,28]. 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 Φ —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 symmetry operations is two. Hence, the symmetry provides the twofold parsimony in the amount of the graphical information necessary for drawing the symbol. It should be emphasized that the aforementioned reasoning does not depend on the specific method of drawing of the symbol. Now, let us quantify the aforementioned parsimony. We denote $m_i(G)$ as the total number of elements of symmetry related to the *i*-th letter of the given alphabet, known in the group theory at the order of the group G [29]. Now, we introduce the symmetry factor of the alphabet denoted μ and defined with Equation (11):

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

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. The value of the symmetry factor μ calculated from the modern English coincides with that established for the Latin script. This finding supports the hypothesis that the time evolution of the Phoenician-rooted scripts stopped with the appearance of the Latin script and justifies the exclusion of the modern English regression line, presented in Figure 6. The straight regression line, shown in Figure 6, spans the time period from 12 BC to 1 BC (which is close to one thousand years), when the time evolution of the Phoenician-rooted scripts stopped.



Figure 5. Analysis of symmetry of the symbols of Phoenician alphabet. (**A**) Symbol $\mathbf{\Phi}$ (qoph) is depicted; mirror axis of symmetry S_2 is shown. (**B**) The entire symbol may be obtained by projection of half a symbol relative to the axis S_2 ; thus, the two-fold parsimony of information is provided. (**C**) Symbol $\mathbf{\otimes}$ (theth) 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 is provided.

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

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

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



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 letters inscribed into rectangles (*r*-scripts).

4. Discussion

Language is a system that we rely on for both interpersonal and intrapersonal communication [30]. It is also one of the core elements of any culture or human civilization [30]. Language is a hallmark that distinguishes human beings from other species [30]. 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, they can become *a muntu*, a "person" [30]. The fundamental question is does the language we speak and write shape the way we think about the world? The problem is extremely perplexing; however, it was hypothesized that the alphabet promotes linear thinking and hierarchical reasoning, such as Aristotelian syllogism [31].

Our investigation quantifies the time evolution of alphabets, rooted in the Phoenician script. 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 only way to investigate 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 as 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 BC for Hebrew and 750 BC for Aramaic, and kept 22 letters of the Phoenician alphabet [30,31].

The reported research presents a kind of phylogenetic approach to computational/ quantitative paleography [32]. Phylogenetics aims to uncover the evolutionary relationships between taxa to obtain an understanding of their evolution. Phylogenetics in a wider sense has three areas: phenetics (numerical taxonomy), cladistics and phylogenetics (in a narrow sense) [32]. Our research presents the narrowest possible approach in the linguistic phylogenetics; namely, we focus exclusively on the time evolution of symmetry of the symbols of alphabets, emerging from the Phoenician one. We neglect the technological and historical context of alphabet evolution (scripts are dependent of the writing materials (stone, wall, wood, ink and paper/papyrus/parchment, etc.)). All of these aspects of the script evolution are neglected in our research, being focused on the evolution of symmetry of symbols only [32].

Why is the symmetry of letters 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 symonymous to symmetry, but it is closely related to it [33–36].

Our investigation addresses the role of symmetry (seen as simplicity) in the emergence of alphabets. It seems intuitively clear that symmetrical letters are easier to draw. How may the effort necessary for their drawing be quantified? We suggest such a measure labeled in our paper the "symmetry factor", which is defined with Equation (11). We demonstrate that symmetrical letters enable the 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 the 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 groups, 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 "Ashurit" or square script, presents an obvious exception from the entire set of Phoenician scripts: quantitative measures of symmetry calculated for these scripts are very different from those established for other Phoenician-rooted scripts. It remains unclear whether the Ashurit script is rooted in Phoenician script [36]. Our findings support the arguments for the point of view, according to which the modern Hebrew/Ashurit script did not emerge from the Phoenician one [36].

5. Conclusions

The 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 computed for the scripts inscribed into a square (abbreviated *s*-scripts) and also for the scripts inscribed into rectangles (abbreviated *r*-scripts) [37]. Actually, H_{SYM} quantifies the average uncertainty to find an element of symmetry within the symbols of the given alphabet (averaged over the entire alphabet). H_{SYM} , calculated for the s-scripts, is decreased with time, whereas H_{SYM} is only slightly time-sensitive for the *r*-scripts. We also calculated the Shannon diversity index (SDI), denoted $D_{SYM}(G)$, for the s- and r-scripts. SDI quantifies the diversity of symmetry groups inherent for the addressed alphabet. We established that SDI increases with time in a monotonic way for both *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 it as follows: $\mu = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{m_i(G)}$, where $m_i(G)$ is 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 the simplification of writing. In other words, the parsimony of graphical information necessary for writing graphemes is increased with time. We conclude that quantification of symmetry of the alphabets could hardly be performed with a single parameter, and a number of quantitative characteristics should be introduced for this purpose, namely, the Shannon measure of symmetry, the Shannon diversity index and the symmetry factor.

We also concluded that the values of the symmetry factor calculated for rectangular and square scripts are close to one another for all of the studied scripts emerging from the Phoenician alphabet, with the pronounced exception of the Hebrew/Ashurit script. Our calculation demonstrated that the time evolution of the Phoenician-rooted scripts stopped with the emergence of the Latin script; indeed, the quantitative parameters of symmetry established for the modern English script are very close to those calculated for the Latin script.

Thus, our study supplies the arguments for the point of view, according to which the modern Hebrew/Ashurit script did not emerge from the Phoenician script [38]. Why is the symmetry of graphemes 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 provides parsimony of graphical information necessary for the drawing of the entire set of letters and (iv) symmetry to a great 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 cannot 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 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. Study of the symmetry of the different patterns (physical or biological) may serve as the Ariadne Thread enabling understanding of their time evolution [41]. And this is also true for the graphical patterns/scripts.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/sym16040465/s1, Single pdf file containing the tables of alphabets.

Author Contributions: Conceptualization, A.G., M.F., S.S. and E.B.; software, A.G. and M.F.; validation A.G. and M.F.; formal analysis, A.G., M.F. and E.B. writing—original draft preparation, E.B.; writing—review and editing, A.G., M.F., S.S. and E.B.; supervision: S.S. and E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors are thankful to Shlomo Dimentman for extremely useful discussions. The authors are thankful to anonymous reviewers for extremely useful reviewing of the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

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" and absence as "0". "1/0" means that the symmetry element is present in the square configuration but not in the rectangular one.

Phoenician Alphabet												
Script	C1	S_1	S_2	S_3	S_4	C4	C ₂	C_4^3				
*	1	1	0	0	0	0	0	0				
٩	1	0	0	0	0	0	0	0				
1	1	0	0	0	0	0	0	0				
Δ	1	0	0	0	0	0	0	0				
Ę	1	0	0	0	0	0	0	0				
Y	1	0	1	0	0	0	0	0				
I	1	1	1	0	0	0	1	0				
目	1	1	1	0	0	0	1	0				
8	1	1	1	1/0	1/0	1/0	1	1/0				
٦	1	0	0	0	0	0	0	0				
k	1	0	0	0	0	0	0	0				
L	1	0	0	0	0	0	0	0				
ግ	1	0	0	0	0	0	0	0				
۲	1	0	0	0	0	0	0	0				
Ŧ	1	1	1	0	0	0	1	0				
0	1	1	1	0	0	0	1	0				
2	1	0	0	0	0	0	0	0				
٣	1	0	0	0	0	0	0	0				
φ	1	0	1	0	0	0	0	0				
٩	1	0	0	0	0	0	0	0				
W	1	0	1	0	0	0	0	0				
Х	1	1	1	1/0	1/0	1/0	1	1/0				

References

- 1. Sampson, G. Writing systems: Methods for recording language. In *The Routledge Handbook of Linguistics;* Allan, K., Ed.; Routledge: New York, NY, USA, 2016; Chapter 4; pp. 47–62.
- 2. Keith, A. What is linguistics? Methods for recording language. In *The Routledge Handbook of Linguistics*; Allan, K., Ed.; Routledge: New York, NY, USA, 2016; Chapter 4; pp. 1–16.
- 3. Gardiner, A.H. The egyptian origin of the semitic alphabet. J. Egypt. Archeol. 1916, 3, 1–16. [CrossRef]
- Pennacchietti, F.A. An alternative hypothesis on the origin of the Greek alphabet, Kervan. *Int. J. Afr. Asiat. Stud.* 2023, 27, 109–121.
 Powell, B.B. *Homer and the Origin of the Greek Alphabet*; Cambridge University Press: Cambridge, MA, USA, 1991; Chapter 1;
- рр. 5–66.

- 6. Hurford, J.R. Evolutionary Linguistics: How Languages and Language Got to Be the Way They Are. In *The Routledge Handbook of Linguistics*; Allan, K., Ed.; Routledge: New York, NY, USA, 2016; Chapter 2; pp. 17–32.
- 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? *Cogn. Sci.* 2007, *31*, 961–987. [CrossRef] [PubMed]
- 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; Ferrara, S., Valério, M., Eds.; Studi Micenei ed Egeo Anatolici: Rome, Italy, 2018; pp. 189–209.
- 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. *Curr. Anthropol.* **2021**, *62*, 669–691. [CrossRef]
- 11. Gibson, E.; Futrell, R.; Piantadosi, S.T.; Dautriche, I.; Mahowald, K.; Bergen, L.; Levy, R. How efficiency shapes human language. *Trends Cogn. Sci.* **2019**, *23*, 389–407. [CrossRef] [PubMed]
- 12. Han, S.J.; Kelly, P.; Winters, J.; Kemp, C. Simplification Is Not Dominant in the Evolution of Chinese Characters. *Open Mind Discov. Cogn. Sci.* **2022**, *6*, 264–279. [CrossRef] [PubMed]
- 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.
- Revesz, P.R. The development and role of symmetry in ancient scripts. In Symmetry: Art and Science | 12th SIS Symmetry Congress [Special Issue]. Symmetry: Art and Science; Viana, V., Nagy, D., Xavier, J., Neiva, A., Ginoulhiac, M., Mateus, L., Varela, P., Eds.; International Society for the Interdisciplinary Study of Symmetry: Budapest, Hungary, 2022; pp. 308–315. Available online: https://repositorio-aberto.up.pt/bitstream/10216/144701/2/588634.pdf (accessed on 12 December 2023).
- 15. Altmann, S.L.; Herzig, P. Point-Group Theory Tables; Oxford University Press: Oxford, UK, 1994.
- 16. Bonfante, J.H.; Bonfante, L.P. The Etruscan Language: An Introduction; Manchester University Press: Manchester, UK, 1983.
- 17. федорова, Е.В. Введение в Латинскую Эпиграфику; Изд-во Московского Университета: Москва, Russian, 1982. (In Russian)
- Ярцева, В.Н. Большой Энциклопедический Словарь Языкознание; Изд-во "Советская Энциклопедия": Москва, Russian, 1989. (In Russian)
- 19. Skolnik, F.; Berenbaum, M. *Encyclopaedia Judaica*, 2nd ed.; Gale, T., Ed.; Macmillan Reference USA: Farmington Hills, MI, USA; Detroit, MI, USA, 2007; Volume 1.
- 20. Fischer, S.R. History of Writing; Reaktion Books: London, UK, 2005; p. 298.
- 21. Jost, L. Entropy and diversity. Oikos 2006, 113, 363–375. [CrossRef]
- 22. Ben-Naim, A. Entropy, Shannon's Measure of Information and Boltzmann's H-Theorem. Entropy 2017, 19, 48. [CrossRef]
- 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, 105123. [CrossRef]
- 24. Radivojević, M. Archaeometallurgy of the Vinča culture: A case study of the site of Belovode in Eastern Serbia. *Hist. Metall.* **2014**, 47, 13–32.
- 25. Radivojević, M.; Rehren, T.; Pernicka, E.; Šljivar, D.; Brauns, M.; Borić, D. On the origins of extractive metallurgy: New evidence from Europe. *J. Archaeol. Sci.* 2010, *37*, 2775–2787. [CrossRef]
- 26. Starović, A. If the Vinča script once really existed who could have written or read it? Doc. Praehist. 2005, 32, 253–260. [CrossRef]
- 27. Rosen, J. Symmetry in Science: An Introduction to the General Theory; Springer: Berlin, Germany, 1995.
- 28. Chatterjee, S.K. Crystallography and the World of Symmetry; Springer: Berlin, Germany, 2008.
- 29. Arfken, G.B.; Weber, H.J. *Mathematical Methods for Physicists*, 5th ed.; A Harcourt Science and Technology Company: San Diego, CA, USA, 2001.
- 30. Pae, H.K. Script E ffects as the Hidden Drive of the Mind, Cognition, and Culture, Literacy Studies. In *Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education;* Springer: Cham, Switzerland, 2020; Volume 21.
- Logan, R.K. The Alphabet Effect: A Media Ecology Understanding of the Making of Western Civilization; Hampton Press: Cresskill, NJ, USA, 2004.
- 32. Rehbein, M.; Sahle, P.; Schassan, T. *Codicology and Palaeography in the Digital Age*; BoD–Books on Demand: Norderstedt, Germany, 2009; Volume 1.
- Wiley, R.W.; Wilson, C.; Rapp, B. The effects of alphabet and expertise on letter perception. *J. Exp. Psychol. Hum. Percept. Perform.* 2016, 42, 1186–1203. [CrossRef] [PubMed]
- 34. Somov, G.U. Interrelations of codes in human semiotic systems. Semiotica 2016, 213, 557–599. [CrossRef]
- 35. Yodogawa, E. Symmetropy, an entropy-like measure of visual symmetry. *Percept. Psychophys.* **1982**, *32*, 230–240. [CrossRef] [PubMed]
- Dry, M.J. Using relational structure to detect symmetry: A Voronoi tessellation based model of symmetry perception. *Acta Psychol.* 2008, 128, 75–90. [CrossRef] [PubMed]
- Bormashenko, E.; Legchenkova, I.; Frenkel, M.; Shoval, S. Shannon (Information) Measures of Symmetry for 1D and 2D Shapes and Patterns. *Appl. Sci.* 2022, *12*, 1127. [CrossRef]
- 38. Hoffman, J.M. In the Beginning. A Short History of Hebrew Language; New York University Press: New York, NY, USA, 2004.
- 39. Bormashenko, E. Entropy, Information, and Symmetry: Ordered is Symmetrical. Entropy 2020, 22, 11. [CrossRef] [PubMed]

- 40. Bormashenko, E. Entropy, Information, and Symmetry; Ordered Is Symmetrical, II: System of Spins in the Magnetic Field. *Entropy* **2020**, *22*, 235. [CrossRef] [PubMed]
- 41. Dyakin, V.V.; Dyakina-Fagnano, N.V. Enigma of Pyramidal Neurons: Chirality-Centric View on Biological Evolution. Congruence to Molecular, Cellular, Physiological, Cognitive, and Psychological Functions. *Symmetry* **2024**, *16*, 355. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.