Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document
Title: Final proposal for encoding the Phoenician script in the UCS
Source: Michael Everson
Status: Individual Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Date: 2004-04-26
Replaces: N1932 (1998-11-23)

This document revises and replaces N1932, which was based on the proposal written by Rick McGowan and published in UTR#3, and the proposal written by me in N1592. It contains the proposal summary.

A. Administrative

1. Title
Final proposal for encoding the Phoenician script in the UCS
2. Requester's name
Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2004-04-26
5. Requester's reference (if applicable)
N1592, N1932
6. Choose one of the following:
6a. This is a complete proposal
Yes.
6b. More information will be provided later
No.

B. Technical – General

1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
Yes.
Proposed name of script
Phoenician.
1b. The proposal is for addition of character(s) to an existing block
No.
1b. Name of the existing block
2. Number of characters in proposal
27.
3. Proposed category (see section II, Character Categories)
Category C
4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)
Level 1.
4b. Is a rationale provided for the choice?
Yes.
4c. If YES, reference
Spacing characters are proposed.
5a. Is a repertoire including character names provided?
Yes.

5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?
Yes.
5c. Are the character shapes attached in a legible form suitable for review?
Yes.
6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson. TrueType.
6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes, see bibliography below.
7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
8. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?
Yes, see below.
9. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.
Yes, see Unicode properties below.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Yes. N1592, N1932
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
No. Phoenician is a simple and well-known historic script used in a wide variety of contexts.
2b. If YES, with whom?
2c. If YES, available relevant documents
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
Scholarly communities researching the Phoenician language; educational communities of all kinds interested in the history of the Latin, Greek, Hebrew, and other alphabets..
4a. The context of use for the proposed characters (type of use; common or rare)
Phoenician script is proposed to unify Proto-Sinaitic/Proto-Canaanite, Punic, Neo-Punic, Phoenician proper, Late Phoenician cursive, Phoenician papyrus, Siloam Hebrew, Hebrew seals, Ammonite, Moabite, Palaeo-Hebrew.
4b. Reference
N2311.
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where?
By scholars and script enthusiasts worldwide.
6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?
No.
6b. If YES, is a rationale provided?
6c. If YES, reference
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Yes, they should be encoded in a single block as presented here.
8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
No.
8b. If YES, is a rationale for its inclusion provided?
8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference
10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
No.
10b. If YES, is a rationale for its inclusion provided?
10c. If YES, reference
11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?
No.
11b. If YES, is a rationale for such use provided?
11c. If YES, reference
12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
No.
12b. If YES, reference
13a. Does the proposal contain characters with any special properties such as control function or similar semantics?
No.
13b. If YES, describe in detail (include attachment if necessary)
14a. Does the proposal contain any Ideographic compatibility character(s)?
No.
14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?
14c. If YES, reference

D. Proposal

The Phoenician alphabet and its successors were widely used over a broad area surrounding the Mediterranean Sea. Phoenician evolved over the period from about the 12th century BCE with some modifications until the 2nd century BCE, with the last neo-Punic inscriptions dating from about the 3rd century CE. Garbini 2001 suggests that while the origins of Phoenician may have been a reform of the Proto-Sinaitic/Canaanite script, it came into its own from the 9th century BCE, when it "became a very elegant script with long, slightly slanting vertical lines, minuscule loops and flat letters." The Phoenician alphabet is a forerunner of the Etruscan, Latin, Greek, Arabic, Hebrew, and Syriac scripts among others, many of which are still in modern use. It has also been suggested that Phoenician is the ultimate source of Kharoshthi and of the Indic scripts descending from Brahmi.

Phoenician is quintessentially illustrative of the historical problem of where to draw lines in an evolutionary tree of continuously changing scripts in use over thousands of years. The twenty-two letters in the Phoenician block may be used, with appropriate font changes, to express Punic, Neo-Punic, Phoenician proper, Late Phoenician cursive, Phoenician papyrus, Siloam Hebrew, Hebrew seals, Ammonite, Moabite, and Palaeo-Hebrew. The historical cut that has been made here considers the line from Phoenician to Punic to represent a single continuous branch of script evolution.

Processing

Phoenician is written from right to left horizontally. Phoenician language inscriptions usually have no space between words; there are sometimes dots between words in later inscriptions (*e.g.* in Moabite inscriptions). Typical fonts for the Phoenician and especially Punic have very exaggerated descenders. These descenders help distinguish the main line of Phoenician evolution toward Punic from the other (*e.g.* Hebrew) branches of the script, where the descenders instead grew shorter over time.

Numerals

Phoenician numerals are built up from four elements in combination; samples of these are shown in Figures 3, 5, 6, 7, 8 and 11. Like the letters, Phoenician numbers are written from right to left: 𐤛𐤚𐤚𐤁𐤁𐤁 means 143 (100 + 20 + 20 + 1 + 1 + 1).

Names

The names used for the characters here are those reconstructed by Theodor Nöldeke in 1904, as given in Powell 1996 (see Figure 1).

Unicode Character Properties

```
10900;PHOENICIAN LETTER ALF;Lo;0;R;;;;;N;;;;;
10901;PHOENICIAN LETTER BET;Lo;0;R;;;;;N;;;;;
10902;PHOENICIAN LETTER GAML;Lo;0;R;;;;;N;;;;;
10903;PHOENICIAN LETTER DELT;Lo;0;R;;;;;N;;;;;
10904;PHOENICIAN LETTER HE;Lo;0;R;;;;;N;;;;;
10905;PHOENICIAN LETTER WAU;Lo;0;R;;;;;N;;;;;
10906;PHOENICIAN LETTER ZAI;Lo;0;R;;;;;N;;;;;
10907;PHOENICIAN LETTER HET;Lo;0;R;;;;;N;;;;;
10908;PHOENICIAN LETTER TET;Lo;0;R;;;;;N;;;;;
10909;PHOENICIAN LETTER YOD;Lo;0;R;;;;;N;;;;;
1090A;PHOENICIAN LETTER KAF;Lo;0;R;;;;;N;;;;;
1090B;PHOENICIAN LETTER LAMD;Lo;0;R;;;;;N;;;;;
1090C;PHOENICIAN LETTER MEM;Lo;0;R;;;;;N;;;;;
1090D;PHOENICIAN LETTER NUN;Lo;0;R;;;;;N;;;;;
1090E;PHOENICIAN LETTER SEMK;Lo;0;R;;;;;N;;;;;
1090F;PHOENICIAN LETTER AIN;Lo;0;R;;;;;N;;;;;
10911;PHOENICIAN LETTER PE;Lo;0;R;;;;;N;;;;;
10912;PHOENICIAN LETTER SADE;Lo;0;R;;;;;N;;;;;
10913;PHOENICIAN LETTER QOF;Lo;0;R;;;;;N;;;;;
10914;PHOENICIAN LETTER ROSH;Lo;0;R;;;;;N;;;;;
10915;PHOENICIAN LETTER SHIN;Lo;0;R;;;;;N;;;;;
10916;PHOENICIAN LETTER TAU;Lo;0;R;;;;;N;;;;;
10917;PHOENICIAN NUMBER ONE;No;0;R;;;;1;N;;;;;
10918;PHOENICIAN NUMBER TEN;No;0;R;;;;10;N;;;;;
10919;PHOENICIAN NUMBER TWENTY;No;0;R;;;;20;N;;;;;
1091A;PHOENICIAN NUMBER ONE HUNDRED;No;0;R;;;;100;N;;;;;
1091F;PHOENICIAN WORD SEPARATOR;Po;0;R;;;;;N;;;;;
```

Bibliography

Faulmann, Carl. 1990 (1880). *Das Buch der Schrift*. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8

Fossey, Charles. 1948. *Notices sur les caractères étrangers anciens et modernes rédigées par un groupe e savants*. Nouvelle édition míse à jour à l'occasion du 21e Congrès des Orientalistes. Paris: Imprimerie Nationale de France.

Garbini, Giovanni. 2001. "The question of the alphabet" in *The Phoenicians*, ed. Sabatino Moscati. I. B. Tauris. ISBN 1-85043-533-2

Healey, John F. 1990. *The early alphabet*. (Reading the past). London: British Museum. ISBN 0-7141-8073-4

Ifrah, Georges. 1998. *The universal history of numbers: from prehistory to the invention of the computer.* London: Harvill Press. ISBN 1-86046-324-X

Imprimerie Nationale. 1990. *Les caractères de l'Imprimerie Nationale*. Paris: Imprimerie Nationale Éditions. ISBN 2-11-081085-8

McCarter, P. Kyle, Jr. 1975. *The antiquity of the Greek alphabet and the early Phoenician scripts*. Missoula, Montana: Scholars Press. (Harvard Semitic Monographs; 9) ISBN 0-89130-066-X

Nöldeke, Theodor. 1904. *Beiträge zur semitischen Sprachwissenschaft*. Strasburg.

Powell, Barry B. 1996. *Homer and the origin of the Greek alphabet*. Cambridge: Cambridge University Press. ISBN 0-521-58907-X .

Robinson, Andrew. 1995. *The story of writing*. London: Thames & Hudson. ISBN 0-500-01665-8

Rumford, James. 2002. *There's a monster in the alphabet*. Boston: Houghton Mifflin. ISBN 0-618-22140-9

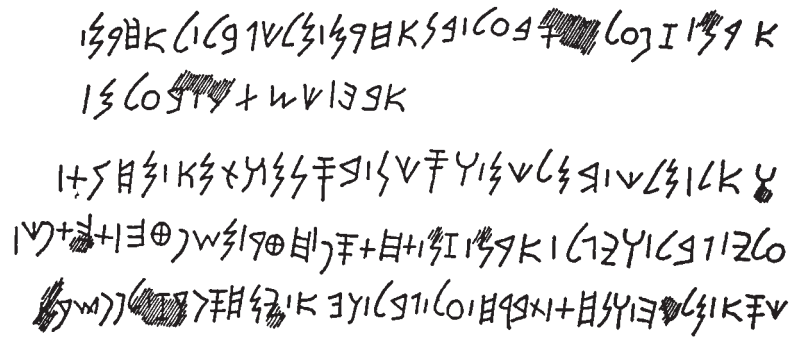
Figures

Table I *The place of early Greek letter forms in the development of Phoenician letter forms*

Phoen. names	PHOENICIAN				GREEK		PHOENICIAN			Greek names
	Shipitbaal I (Byblos) (= *DR* No. 7) *c.* 900	Cyprus Stele (= *DR* No. 30) *c.* 900–875	Nora Stone (Sardinia) (= *DR* No.40) 9th cent.	Kilamuva (Zincirli) (= *DR* No. 24) *c.* 825	Dipylon jug (Athens) *c.* 740	Lefkandi, Pithekoussai *c.* 750–25	Limassol (Cyprus) (= *DR* No. 31) *c.* 750–25	Karatepe (= *DR* No. 26) *c.* 720	Ipssambul *c.* 590	
ʾalf										alpha
bēt										bēta
gaml										gamma
delt										delta
hē										ei
wau										[wau]
zai										zēta
ḥēt										[h]ēta
ṭēt										thēta
yōd										iōta
kaf										kappa
lamd										lambda
mēm										mū
nūn										nū
semk										xei
ʿain										ou
pē										pei
ṣādē										san
qōf										qoppa
rōš										rhō
šin										sigma
tau										tau

All signs are drawn from right to left.
Phoenician forms are based on Friedrich–Röllig, 1970: end table.

Figure 1. Table of Phoenician and Greek letterforms from Powell 1996.
The character names taken from Theodor Nöldeke's reconstruction are shown in the first column.



12 Phoenician inscription of Ahiram. From Byblos, late 11th century BC.

Figure 2. Text sample from Healey 1990.
Note the use of the PHOENICIAN WORD SEPARATOR with a short vertical glyph.

corps 18 Trois inscriptions de Larnaca

Figure 3. Sample from Imprimerie Nationale 1990:161. Note the use of digits (6 and 1 in line 1, and 2 in line 3) and of the PHOENICIAN WORD SEPARATOR with a dotted glyph in the second passage.

PLATE II

COLONIAL PHOENICIAN INSCRIPTIONS, BEFORE 800 B.C.

PLATE III

COLONIAL PHOENICIAN INSCRIPTIONS, EIGHTH CENTURY B.C.

Figure 4. Plates from McCarter 1975 showing different glyph variants of Phoenician script.

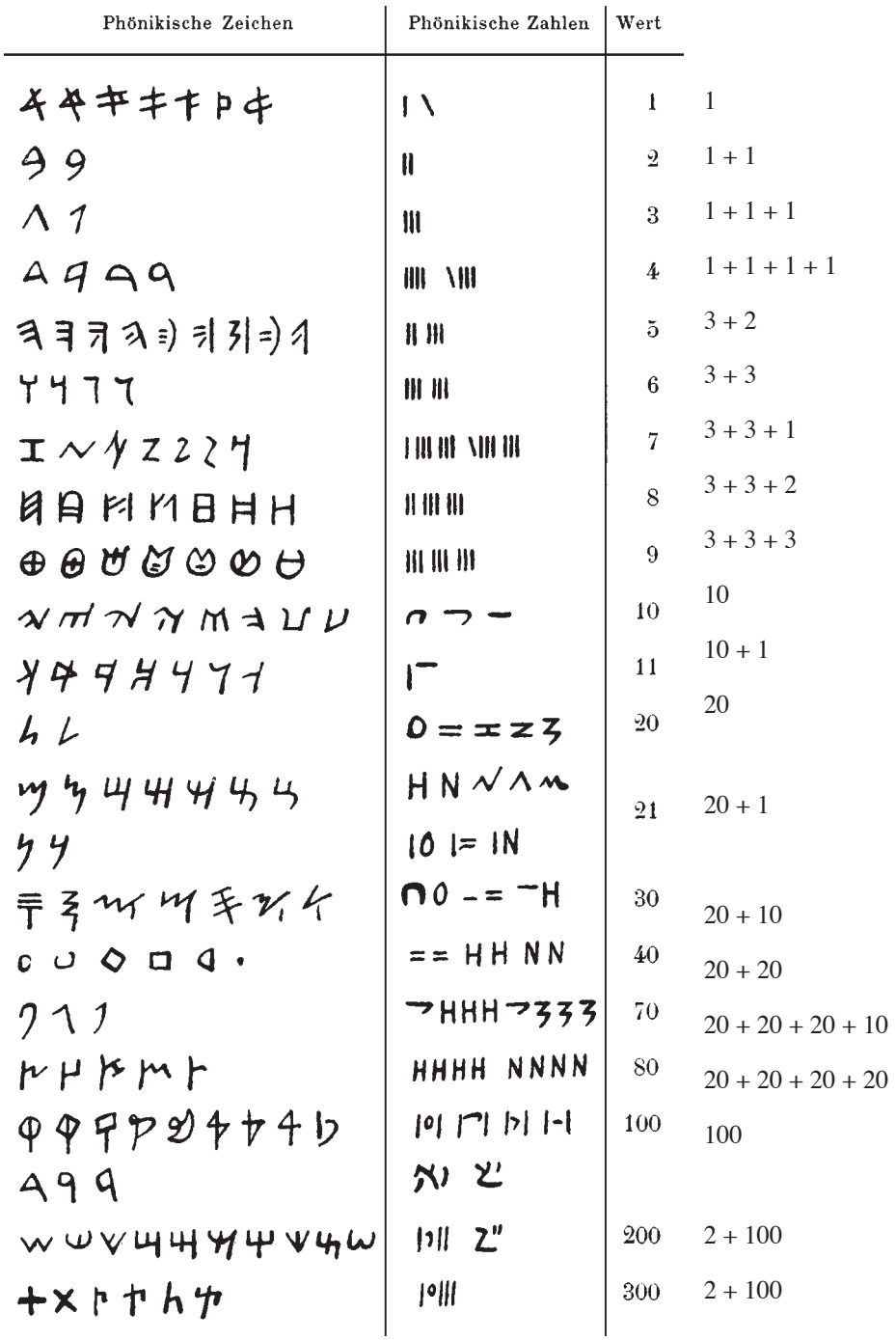


Figure 5. Sample from Faulmann 1880 showing glyph variants for Phoenician letters and numbers.

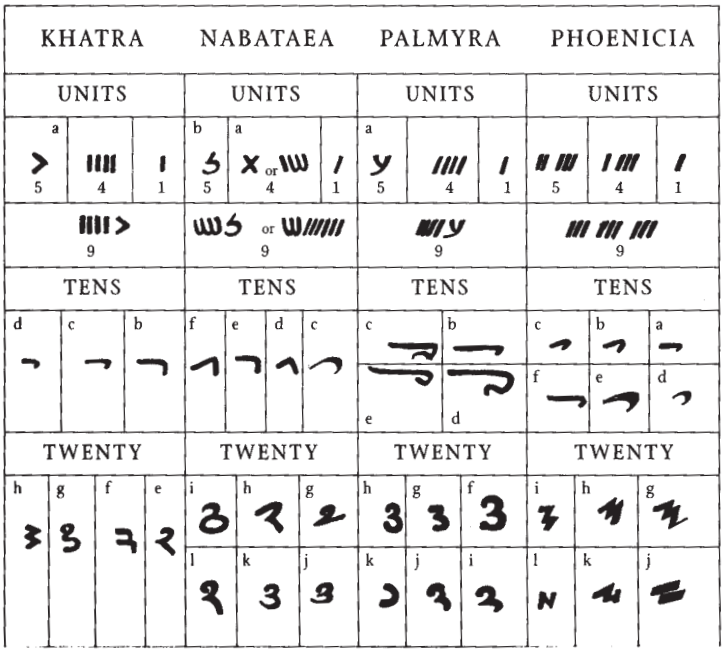


Figure 6. Sample from Ifrah 1998 showing the Phoenician numbers ONE. TEN, and TWENTY.

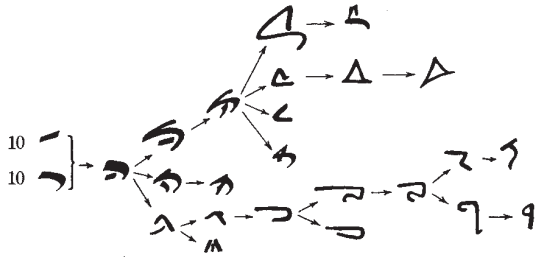


FIG. 18.6. *Origin and development of the figure 100. All these signs derive from placing two variants of the sign for 10 one above the other. This multiplicative combination has a kind of additional superscript to avoid confusing it with the sign for 20, and produced widely different graphical representations of the number 100.*

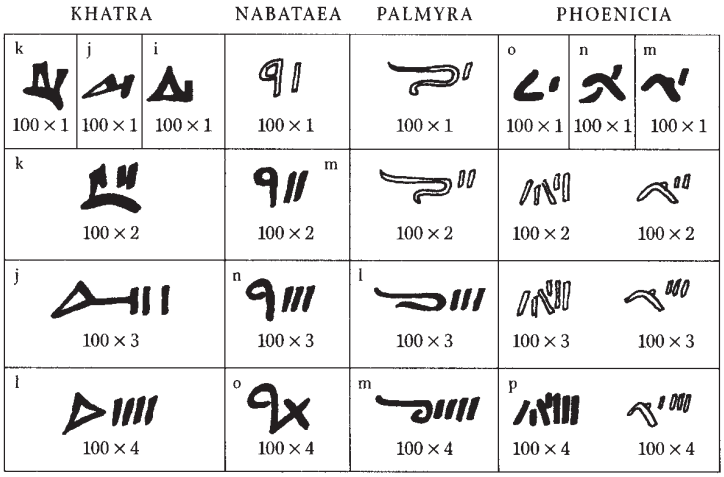


FIG. 18.7. *Semitic representations of the number 100. Attested examples are given in solid lines; reconstructed examples in outline. For sources, see list of references in Fig. 18.2 and 18.5.*

Figure 7. Sample from Ifrah 1998 showing the Phoenician number ONE HUNDRED.

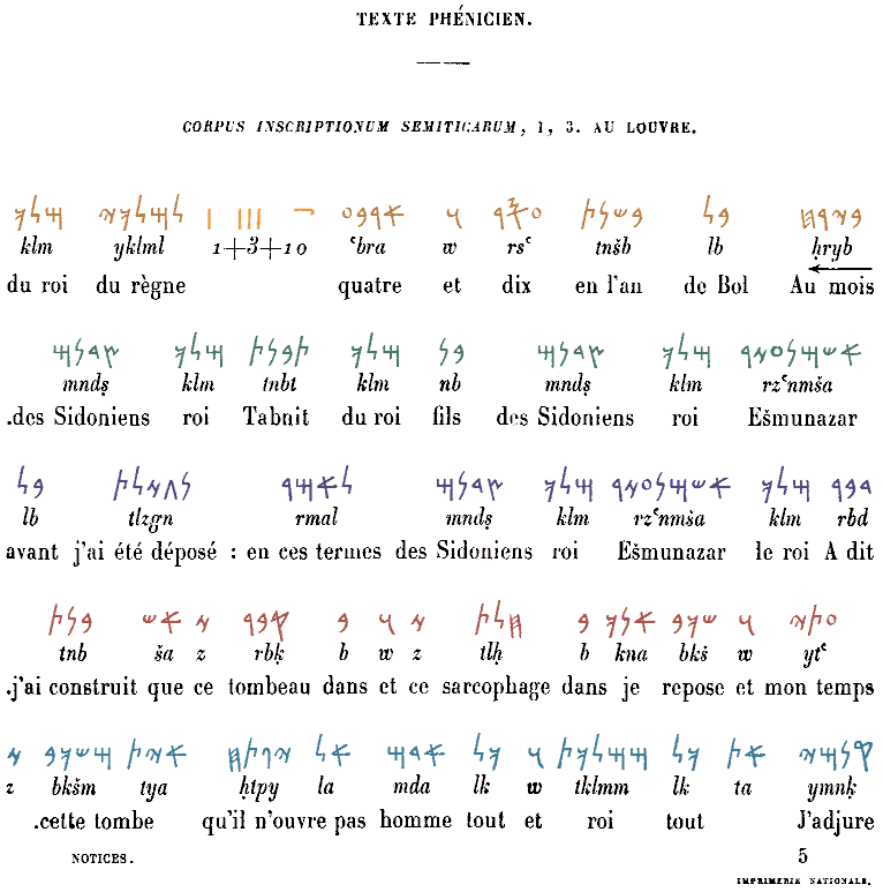


Figure 8. Phoenician text sample from Fossey 1948. The text is from the Sarcophagus of Eshmunazar which is now in the Louvre. Compare the coloured lines with those in Figure 11.

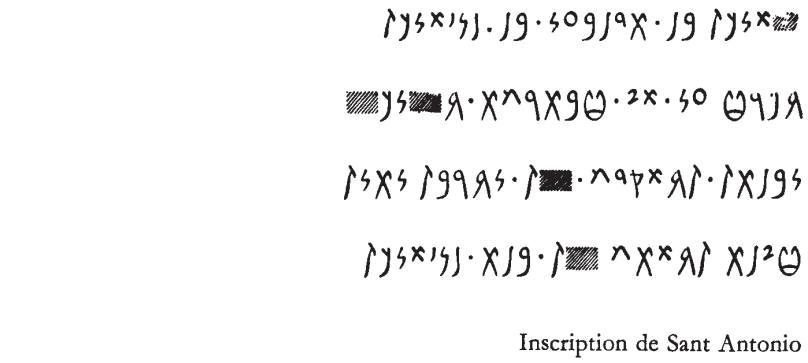


Figure 9. Punic text sample from the Imprimerie Nationale 1990, showing a Punic style of writing. Note the use of the PHOENICIAN WORD SEPARATOR with a dotted glyph here.

Figure 10. Phoenician font samples. Sidon (Michael Everson); Eshmoon (Salim G. Khalaf), Phoenician Moabite (David Myriad Rosenbaum), and Phtem (Maroun Kassab)

Figure 11. Sample from Imprimerie Nationale 1990.
The text is from the Sarcophagus of Eshmunazar which is now in the Louvre. The Phoenician numbers ONE and TEN can be seen in the first line. Compare the coloured lines with those in Figure 8.
A translation of the text of this inscription can be found at **www.shsu.edu/~his_ncp/Eshmun.html**.
The reference glyphs for the code chart are based on this font style; the sarcophagus is shown below.

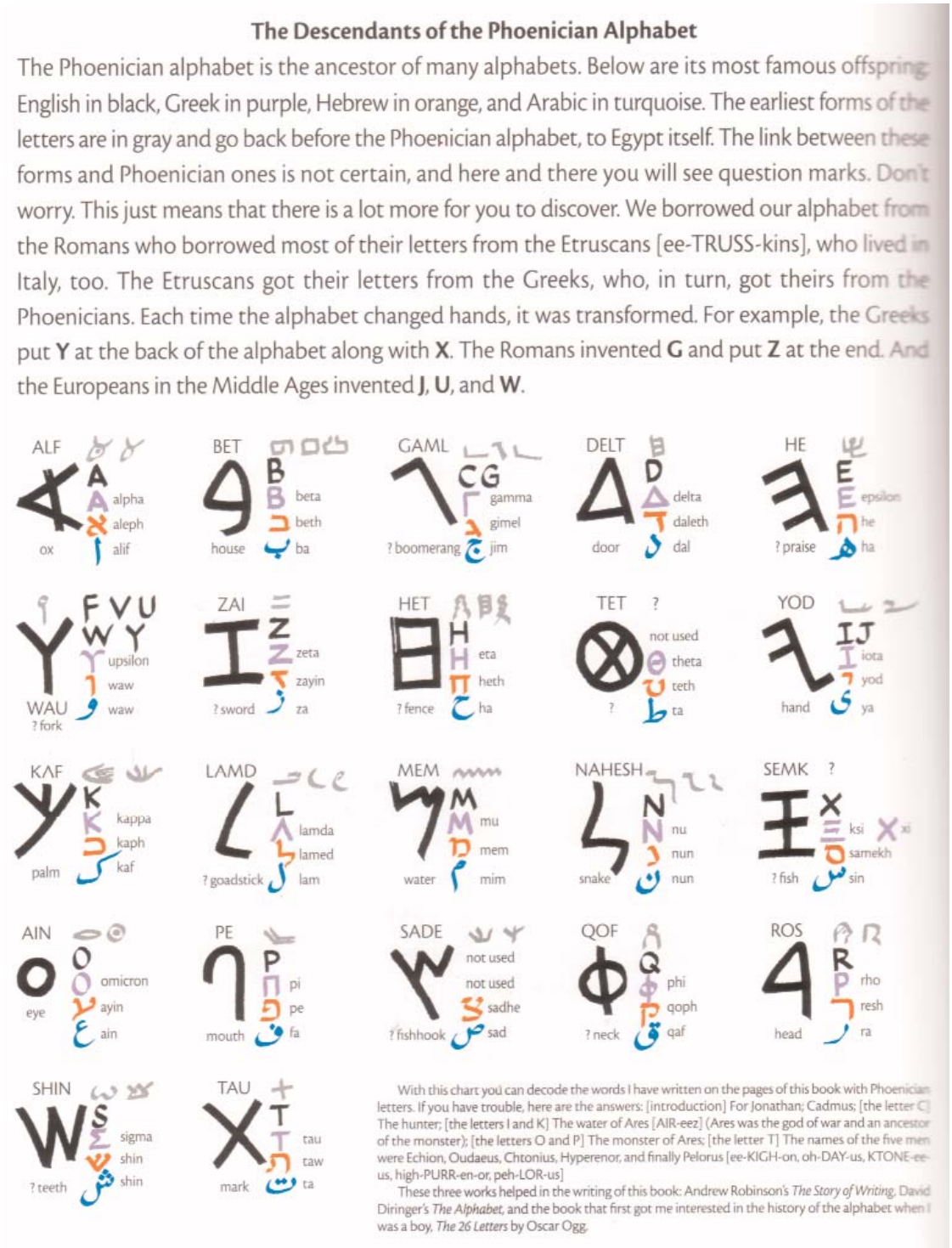


Figure 12. Sample from Rumford 2002. This delightful children's book is an example of non-scholarly yet educational use of Phoenician script in the context of the history of our alphabet. The character names Rumford uses are based on Theodor Nöldeke's reconstruction, except that he uses NAHESH 'snake' rather than NUN 'fish'.

Figure 13. A fragment of the Septuagint dated between 50 BCE and 50 CE. The fragments is part of the "Nahal Hever Minor Prophets" collection, containing fragments of Jonah, Micah, Nahum, Habakkuk, Zephaniah and Zechariah found in the Nahal Hever cave, south of Qumran. The Tetragrammaton in Phoenician script is indicated with the large black arrow; the rest of the text is Greek.

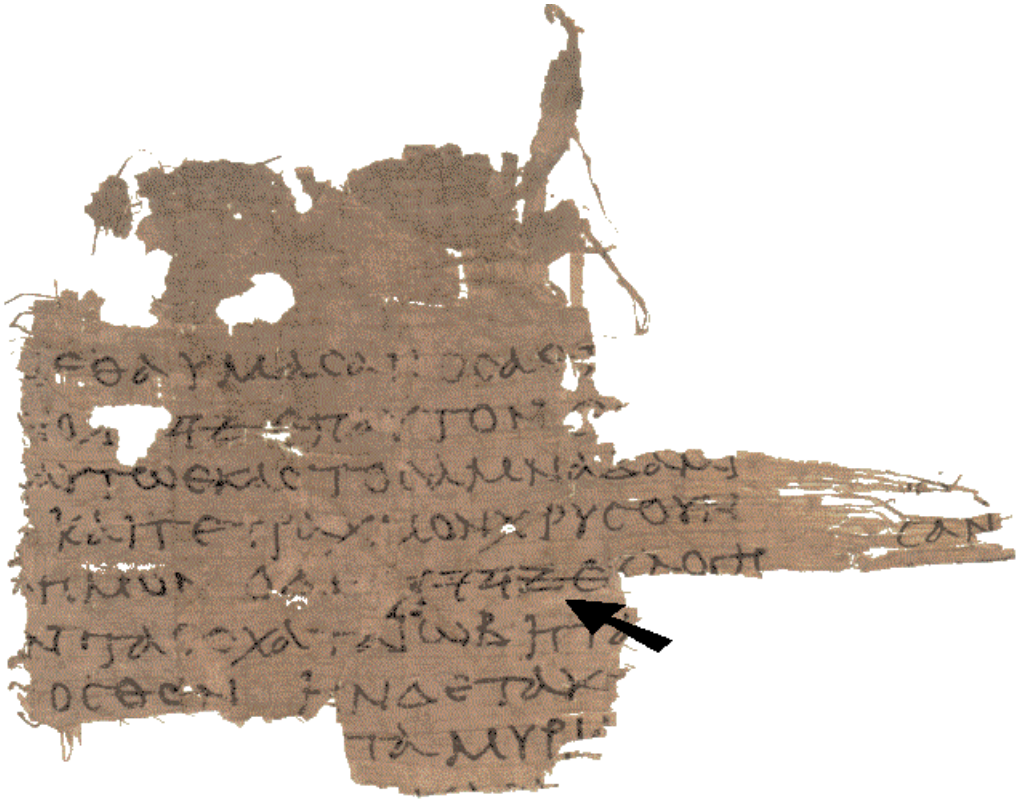


Figure 14. A fragment of Job 42, again containing the Tetragrammaton in Phoenician script alongside Greek text. Apparently no copies of the Septuagint dated before the mid-2nd century CE substitutes the Tetragrammaton with Κυριος 'LORD'. Sample from **www.eliyah.com/lxx.html**

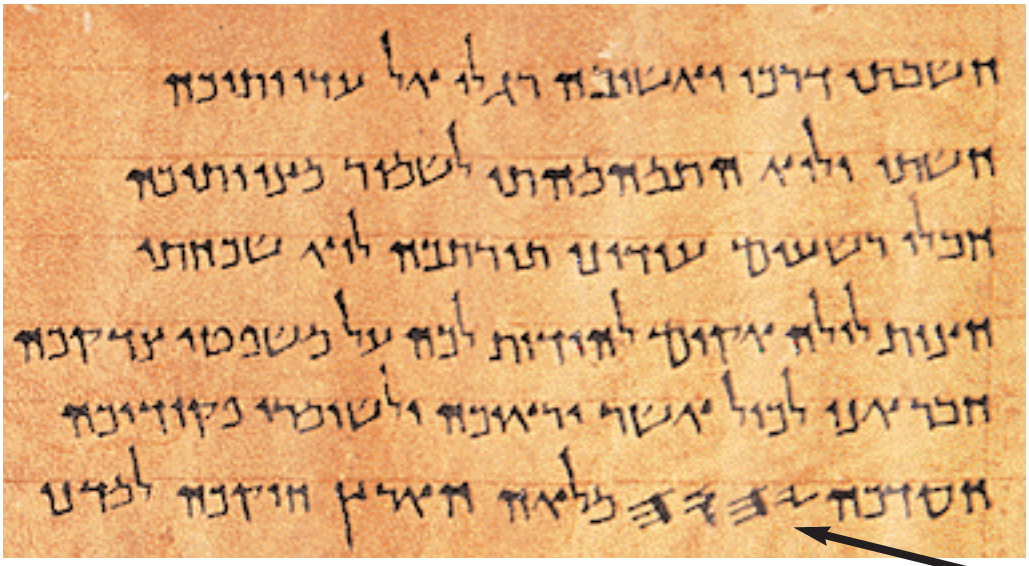


Figure 15. A text from Qumran, containing the Tetragrammaton in Phoenician script (Palaeo-Hebrew variant) alongside Hebrew text.

(PHÉNICIEN, GREC ET NÉO-PUNIQUE.)

VALEUR.	PHÉNICIEN ARCHAÏQUE.	PHÉNICIEN ANCIEN.	NOM.	GREC.	NOM.	PHÉNICIEN RÉCENT.	NÉO-PUNIQUE.
ʾ			aleph	A	alfa		
b			bet	B	bēta		
g			gimel	Γ	gamma		
d			dalet	Δ	delta		
h			hé	E	epsilon		
w			waw	Y F	upsilon digamma		
z			zaïn	Z	zēta		
ḥ			ḥet	H	ēta		
ṭ			ṭet	Θ	ṯēta		
y (i)			yod	I	iōta		
k			kaf	K	kappa		
l			lamed	Λ	lambda		
m			mem	M	mu		
n			nun	N	nu		
s			samek	Ξ	xi		
ʿ			ʿaïn	O	omicron		
p (ph)			pé	Π	pi		
ṣ			ṣadé	M	san		
ḳ			ḳof	ϙ	ḳoppa		
r			reš	P	rhō		
š			šin	Σ	sigma		
t			taw	T	tau		

Figure 16. Table of Archaic Phoenician, Old Phoenician, Greek, Late Phoenician, and cursive Neo-Punic letterforms from Fossey 1948.

TABLE XX - Row 109: PHOENICIAN

	1090	1091
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
A		
B		
C		
D		
E		
F		

G = 00
P = 01

TABLE XX - Row 109: PHOENICIAN

hex	Name
00	PHOENICIAN LETTER ALF
01	PHOENICIAN LETTER BET
02	PHOENICIAN LETTER GAML
03	PHOENICIAN LETTER DELT
04	PHOENICIAN LETTER HE
05	PHOENICIAN LETTER WAU
06	PHOENICIAN LETTER ZAI
07	PHOENICIAN LETTER HET
08	PHOENICIAN LETTER TET
09	PHOENICIAN LETTER YOD
0A	PHOENICIAN LETTER KAF
0B	PHOENICIAN LETTER LAMD
0C	PHOENICIAN LETTER MEM
0D	PHOENICIAN LETTER NUN
0E	PHOENICIAN LETTER SEMK
0F	PHOENICIAN LETTER AIN
10	PHOENICIAN LETTER PE
11	PHOENICIAN LETTER SADE
12	PHOENICIAN LETTER QOF
13	PHOENICIAN LETTER ROSH
14	PHOENICIAN LETTER SHIN
15	PHOENICIAN LETTER TAU
16	PHOENICIAN NUMERAL ONE
17	PHOENICIAN NUMERAL TEN
18	PHOENICIAN NUMERAL TWENTY
19	PHOENICIAN NUMERAL ONE HUNDRED
1A	(This position shall not be used)
1B	(This position shall not be used)
1C	(This position shall not be used)
1D	(This position shall not be used)
1E	(This position shall not be used)
1F	PHOENICIAN WORD SEPARATOR

hex	Name