

## Unicode Technical Report #3 Exploratory Proposals

*Review period closes August 15, 1993  
Another draft will subsequently be issued for review*

### **Introduction**

This Technical Report is comprised of several exploratory proposals that the Unicode Technical Committee wishes to present for their first public review and commentary. These proposals have been generated from the committee's current knowledge about the scripts in question. Most of them are believed to be reasonable technical solutions for encoding of particular scripts, as far as can be ascertained at this time. However, many of them are known to be incomplete or be possessed of significant unresolved issues. The major unresolved issues are discussed in each proposal.

Technical inaccuracies and ambiguities are to be expected in a work of this nature, and most probably abound in these proposals. The work involves conjecture, relies on scanty information, and often requires re-interpretation as new information becomes available. The committee is *not* strongly committed to these proposals as they stand, and further information is being actively sought. Suggestions for improvement by way of additional symbols, further technical requirements, changes in the script model, refinements to the block introductions, or any other information can be mailed to the Scripts Subcommittee at the Unicode, Inc. address. The committee especially wishes to invite active participation and feedback from the communities which these proposals are designed to serve.

In these exploratory proposals, it is often mentioned that "sufficient information is not available" for some particular aspect of the script under discussion. This does not refer to the availability of information in an absolute sense, rather that the committee has not yet been able to obtain sufficient information for its archives.

## Syriac

The Syriac script is a later descendent of the Aramaic script. The earliest known Syriac inscriptions are dated about 6 AD from near the town of Edessa to write the Aramaic dialect that became Syriac. The Syriac script really represents a family of three closely related writing styles called *Estrangela*, *Nestorian*, and *Serta* (the latter is also called *Jacobite*). The earliest form that became distinguished from Aramaic itself is *Estrangela*, developed about the 5th century AD. It was used extensively from the earliest times to record various Christian scriptures. The Syriac script is still in modern use. According to Healey (1990):

“Syriac speaking communities have survived in large numbers in the area around the point where the borders of Syria, Turkey, and Iraq meet, and there are also *emigré* communities in Europe and the United States. Books, magazines and newspapers are still produced in the Syriac scripts.”

The Syriac scripts are generally cursive or semi-cursive, with some letters joining regularly to others and sometimes changing shape in a manner similar to the Arabic script. Vowel signs are known to exist, but available sources do not discuss them.

**Issues:** The vowel signs at least must be added to complete the Syriac proposal. There seem to be at least two different non-spacing vowel systems: one is attributed to Jacob of Edessa and utilizes small letters written above or below others to indicate following vowels; the other is an older dotting system.

The chart shows in parallel the Mandaic alphabet (which includes the extra letter *e* at the end). It is not clear whether Mandaic should be unified with the Syriac block or not; it might be better encoded using the Aramaic block, or encoded separately.

Note that this order differs from the Early Phoenician and Aramic orders. It is not known whether *waw* in particular should come at the end, or at its place here.

### Some Sources

Healey, John F. *The Early Alphabet*.  
Diringer, David. *Writing*.

Mandaic

|   |   |   |  |  |  |  |   |   |
|---|---|---|--|--|--|--|---|---|
| 0 | ܐ | ܒ |  |  |  |  | ܘ | ܙ |
| 1 | ܘ | ܝ |  |  |  |  | ܠ | ܡ |
| 2 | ܚ | ܘ |  |  |  |  | ܢ | ܐ |
| 3 | ܓ | ܕ |  |  |  |  | ܦ | ܩ |
| 4 | ܚ | ܕ |  |  |  |  | ܠ | ܡ |
| 5 | ܘ | ܝ |  |  |  |  | ܠ | ܡ |
| 6 | ܐ |   |  |  |  |  | ܠ | ܡ |
| 7 | ܚ |   |  |  |  |  | ܠ |   |
| 8 | ܚ |   |  |  |  |  | ܠ |   |
| 9 | ܘ |   |  |  |  |  | ܠ |   |
| A | ܘ |   |  |  |  |  | ܠ |   |
| B | ܘ |   |  |  |  |  | ܠ |   |
| C | ܘ |   |  |  |  |  | ܠ |   |
| D | ܘ |   |  |  |  |  | ܠ |   |
| E | ܘ |   |  |  |  |  | ܠ |   |
| F | ܘ |   |  |  |  |  | ܠ |   |

Syriac Names List, draft 92/10/29

- 00 SYRIAC LETTER ALAP
- 01 SYRIAC LETTER BET
- 02 SYRIAC LETTER GAMAL
- 03 SYRIAC LETTER DALAT
- 04 SYRIAC LETTER HE
- 05 SYRIAC LETTER WAW
- 06 SYRIAC LETTER ZAYN
- 07 SYRIAC LETTER HET
- 08 SYRIAC LETTER TET
- 09 SYRIAC LETTER YO
- 0A SYRIAC LETTER KAP
- 0B SYRIAC LETTER LAMAD
- 0C SYRIAC LETTER MIM
- 0D SYRIAC LETTER NUN
- 0E SYRIAC LETTER SEMKAT
- 0F SYRIAC LETTER E
  
- 10 SYRIAC LETTER PE
- 11 SYRIAC LETTER SADE
- 12 SYRIAC LETTER QOP
- 13 SYRIAC LETTER RES
- 14 SYRIAC LETTER SIN
- 15 SYRIAC LETTER TAW

## About the Epigraphic Blocks

### Semitic Alphabets

In these exploratory proposals, we distinguish two major “Early Semitic Alphabet” blocks, **Phoenician** and **Early Aramaic**, which are divided based on what may be termed “significant” differences in the shapes of various letters. Admittedly, this is a highly subjective choice. This arrangement makes two decisive cuts in a historical continuum covering several thousand years of middle-eastern history. The first cut is at approximately the point where several scripts leading eventually to the Aramaic and Hebrew branches began to be quite differentiated in their appearances from the branch that led to Punic. The second cut is at the point where the Aramaic/Hebrew branch began to noticeably split apart into the various lines that led to the Greek, Etruscan, and Latin branches on the one hand, and the Syriac, Arabic, and Hebrew branches on the other.

The alphabet encoded in the **Early Phoenician** block represents Phoenician as it stabilized by about 1100-1050 BC, as well as several early scripts that are quite closely related, though they are used to write a number of languages. The Phoenician block may be used, with appropriate font changes, to express Early Phoenician, Moabite, Early Hebrew, the earliest Early Aramaic, and Canaanite or Proto-Sinaitic scripts. It is also recommended for use to express Later Phoenician and **Punic**, which represent the main line of Phoenician evolution as a distinct script.

### Later Branches of the Phoenician Alphabet

For encoding of **Late Aramaic** (especially papyri), **Palmyrene**, and **Nabataean** the Early Aramaic block should be used. The dividing line is relatively fuzzy, but in general a decision of which block to use can be made on the language, or when necessary on the general appearance of the script. The Unicode blocks are based rather roughly on “significant” differences in at least 12 letters (out of 22), including most obviously the letters transcribed as *A(aleph)*, *B*, *H-underdot*, *T-underdot*, *Y*, *S*, and *R*. (A reasonable comparative source chart is contained in Healey’s *The Early Alphabet*, fig. 15; the two blocks are divided approximately between the fourth and fifth of eight columns.)

## Related Historical Script Blocks

- **South Arabian** and its descendents used for the **Lihyanite**, **Safaitic**, and **Thamudic** languages are encoded in the South Arabian block.
- The **Syriac** scripts (**Serta**, **Estrangela**, and **Nestorian** and their immediate precursors such as **Mandaic**) are encoded in a Syriac block and treated as font differences from a prototypical Syriac script. (Mandaic shapes are also shown in the Syriac block.) Varieties of Syriac are in modern use.
- **Etruscan** and **Oscan** are encoded in the Etruscan block.

## Scripts Not Considered for Encoding

- **Lydian**, **Lycian**, **Sidetic**, **Carian** are not currently being considered for encoding. Information on the repertoire for the first two is available, but other significant information is lacking for all of them. They may eventually be encoded separately, or mapped onto other scripts.

## Future Directions

In the future, this epigraphic introduction may be expanded to include further discussions of epigraphic scripts and families of scripts.

## Some Sources

Healey, John F. *The Early Alphabet*.

Cross, Frank Moore. *The Invention and Development of the Alphabet*.

Encyclopaedia Britannica, Articles on: *Anatolian Languages, Ancient Epigraphic Remains, Alphabets, Luwian, Lycian alphabet, Lycian language, Lydian language*.

Rev 92/11/25

SARGON HASSO  
1649 Pebble Beach Circle  
Elgin, IL 60123

Peter Jensen  
2317 W. Farwell  
Chicago, Ill. 60645



Please find attached:  
✓ Syriac exploratory  
proposals  
✓ Epigraphic bible intro  
for Tech Report #3  
✓ Tech Report #3  
Introduction