

February 3, 1993

Rick McGowan
NeXT, Inc.
900 Chesapeake Drive
Redwood City, CA 94063

Dear Rick:

I want to thank you for the opportunity to contribute to the development of the "Syriac" standard. Enclosed please find a paper that I wrote for the Journal of the Assyrian Academic Society (JAAS). I thought it appropriate to send it to you, since it presents a solution to the very problem you are dealing with, namely, creating a standard code for Assyrian characters and symbols. This paper has been submitted for publication in the next issue of JAAS. Please do not distribute or publish this advance copy I have sent you.

Appendix A of the paper contains SACII, the *Standard Assyrian Code for Information Interchange* (I was tempted to call it ASCII. . .). With SACII, 99% of all Assyrian documents and manuscripts can be reproduced -- whether they are from 500 A.D. or from 1993 A.D.. Appendix A specifies each character, symbol, or diacritical mark, its shape, and its name. Since SACII is a comprehensive encapsulation of the writing elements of the Assyrian language, it should become the working proposal.

The preferred spelling of *Estrangelo* is with an o at the end, not an a. The names "Jacobite" and "Nestorian" refer to religious denominations within the Assyrian nation, and are unsuitable as font names. "Nestorian" is unacceptable to those Assyrians who are of the Assyrian Church of the East because 1) The Assyrian Church of the East predates Nestoris by 400 years, and is not a follower of Nestoris, as the name implies, and 2) there is a heresy stigma attached to Nestoris's name (Nestoris was condemned as a heretic for stating that Mary is the mother of Jesus and not the mother of God. At the time the Church supported him because the church held (and still holds) this doctrinal position, hence the Church was accused of "Nestorianism"). Furthermore, the ~~Roman~~ Catholic Assyrians ("Chaldeans") also use the eastern font, and they are certainly not "Nestorians". "Jacobite" refers to those Assyrians who followed Bishop Jacob after he split from the Assyrian Church of the East in 450 A.D.; The proper name for this church is the Syrian Orthodox Church. **Eastern Assyrian Font** should be used in place of "Nestorian", and **Western Assyrian Font** should be used in place of "Jacobite" or "Serto" (Serto means tied together, or connected, and is misleading, because *all* Assyrian fonts are connected). Eastern and Western can be used in short; they accurately reflect who uses these fonts because, in general, the western Assyrians (in Syria and Lebanon) use all three fonts, while the eastern Assyrians (in Iraq, Iran, and Turkey) use *Estrangelo* and Eastern only.

by the Byzantines
He established
the Monophysite
Church.

Nestorians

BAR
ADDAI

Two issues were raised in the exploratory proposal: the order of the letters and Mandaic. The order of the letters is correct as you have it -- *Wow* is in its correct place. *Wow* should *not* be placed at the end because *Wow* is, by virtue of its position in the Assyrian alphabet (the sixth letter) also the number 6. Assyrian letters are also used as numbers and have ordinal values (if you need more information on this, please let me know). As for Mandaic, when we designed SACII we had Mandaic in the back of our mind. Hence, we left code 87 reserved for the extra letter in Mandaic. It then becomes a simple matter of treating Mandaic as a font, on par with Eastern, Western, or *Estrangelo*.

I now turn to an issue of great significance to the Assyrian people; namely, the use of the word "Syriac" instead of Assyrian. Let me give you a little background. The Assyrian people constitute, in all probability, the oldest surviving nation in the world. The earliest recorded Assyrian is king Sargon I, circa 2400 B.C., and he lists 173 kings before him. Up to about 700 B.C., during the Neo-Assyrian empire (the third, last, and greatest of the Assyrian empires), Assyrians spoke Akkadian, a semitic language, and wrote in cuneiform. But by about 700 B.C. Aramaic had replaced Akkadian (a process that began around 1100 B.C.). The Empire fell in 612 B.C., but Assyrians continued living in their homeland (present day North Iraq), and continued speaking Aramaic, which by 50 A.D. had evolved into "Syriac". Aramaic, as you know, was the *lingua franca* of the Middle East, and was supplanted by Arabic only in the sixth and seventh centuries A.D.. Assyrians, though, preserved their language (which has a distinct Akkadian influence, by the way), and to this day continue to speak, read, and write it. "Syriac" is the oldest continuously spoken language in the world, and only Assyrians are its guardians.

Which brings me to my point. Assyrians do not call their language "Syriac" -- they call it Assyrian, after their national identity. There are two to three million Assyrians in the world; only they use "Syriac". All modern literature in "Syriac" is produced by Assyrians. Given that the users of "Syriac" are Assyrians, and they produce all the literature in "Syriac", then by the *right of national self determination* Assyrians wish their language and their national identity to have the same name -- Assyrian.

When Ben-Yehuda reconstructed the Hebrew language in 1871 in Germany, he borrowed heavily from Aramaic and Syriac (70% of modern Hebrew is derived from Aramaic/Syriac). Yet the language is called Hebrew, after the people who speak it, and rightly so. Assyrians are asking for the same -- no more or no less.

I know that there will be opposition to this from non-Assyrian Syriac scholars. With all due respect, it is none of their business. This is a matter of national self determination. Their scholarship will in no way be affected by the name of the language.

The proper names to use are: **Modern Assyrian** (for the language used by today's Assyrians), **Classical Assyrian** (for the language used in the Church liturgy, which is the same language as

it was in 400 A.D.), and **Ancient Assyrian** (Akkadian). Since we are dealing primarily with today's language, **Modern Assyrian** should be used.

My question, therefore, is this: how can I insure that **Modern Assyrian** is used in the Unicode standard, and not "Syriac"? Your help in this will be greatly appreciated. I ask this because I love my people and my language, and I wish to see them both thrive and prosper, especially in the face of racist Arabization policies imposed on Assyrians by Arab governments (remember, Assyrians are not Arabs and are Christians -- a double whammy in the Middle East).

In the meantime, please look over Appendix A and get back to me with your thoughts. I offer my fullest services to the Unicode committee for developing the Assyrian standard.

Yours sincerely

Peter Jasim, President

Assyrian S+0000 - S+0000

The Assyrian script ("Syriac") is used for writing the Assyrian language, which includes the Eastern Assyrian font ("Nestorian"), the Western Assyrian font ("Serto" or "Jacobite"), and the *Estrangelo* font. It is also used for writing Mandaic.

Assyrian is written from right to left, and is ^{predominantly} cursive, even in its printed form. ^{Few} The same letter may be written in different forms, depending on how it joins to its neighbors. Vowels are written as *zaw'e* applied to consonantal base letters, but are usually omitted in normal writing.

Mandaic. Mandaic is written with the same script, with an additional, 23rd letter; this extra letter is given the independent code S+0022.

Punctuation. Most punctuation marks used with Assyrian are not given independent codes (they are unified with the Latin, Arabic, and Hebrew punctuation), except for the few cases where the mark has a unique form in Assyrian.

Encoding Principles. The alphabet of Assyrian is well defined. Each letter receives only one Unicode character value, regardless of the number of contextual shapes it may exhibit in text (this, indeed, is the only difference between Eastern Assyrian, Western Assyrian, *Estrangelo*, and Mandaic). The graphic form (glyph) shown in the Unicode character chart for an Assyrian letter is that of free-form *Estrangelo*.

Paroshe. *Paroshe* (diacritical points) are marks that indicate vowels (*zaw'e*), cantillation marks, and other modifications of consonantal letters. The occurrence of a character in the *Paroshe* range, and its depiction in relation to a dashed circle, constitute an assertion that this character is intended to be applied via some process to the consonantal letter that precedes it in the text stream, the base character. General rules for applying non-spacing marks are given in the Generic Diacritical Mark block description section. The Unicode standard does not specify a sequence order in case of multiple marks applied to the same Assyrian base character, since there is no possible ambiguity of interpretation. The Assyrian script contains a rich set of diacritical marks, which reflects its development over the course of its long history (it is the oldest spoken language in the world).

Encoding Structure. The Assyrian block is divided into the following ranges:

S+0700-S+0022	Assyrian letters	④
S+0723-S+07??	Assyrian punctuation	①
S+07??-S+07??	Non-spacing characters, Assyrian vowels	③ → ②
S+07??-S+07??	Non-spacing characters, above the line	②
S+07??-S+07??	Non-spacing characters, on and above the line	
S+07??-S+07??	Non-spacing characters, below the line	
S+07??-S+07??	Non-spacing characters, on and below the line	
S+07??-S+07??	Non-spacing characters, additional	③

?

ORDER

Eastern Assyrian Font should be used in place of "Nestorian", and Western Assyrian Font should be used in place of "Jacobite" or "Serto" (Serto means tied together, or connected, and is misleading, because *all* Assyrian fonts are connected). Eastern and Western can be used in short; they accurately reflect who uses these fonts because, in general, the western Assyrians (in Syria and Lebanon) use all three fonts, while the eastern Assyrians (in Iraq, Iran, and Turkey) use *Estrangelo* and Eastern only.

redundant!

notes:

names of various accents is ~~language~~ only suggestive since the same form is used for different purposes & different ^{connotation} meanings.

- Accent
- diacritical signs - marks

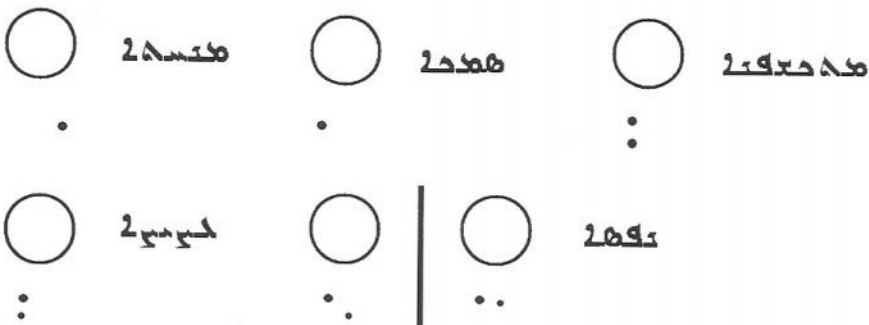
Points above the line



Points on the line



Points below the line



Points on and above the line



Points on and below the line



REVISED

October 11, 1993

Mr. Rick McGowan
NeXt, Inc.,
900 Chesapeake Drive
Redwood City, CA 94063

Dear Rick:

A while back I had spoken with you on the phone regarding the inclusion of the Syriac script in the Unicode standard. Shortly thereafter, Peter Jasim had some correspondence with you regarding the same matter. You indicated to me then that no one had submitted any proposal yet. Therefore, we are submitting this proposal to the Unicode Consortium for thorough examination and possible inclusion in the Unicode standard. Please feel free to contact me for any question and/or clarification that you, or the committee in charge, might have.

Sincerely,

Sargon Hasso, Independent Consultant

P.S.

Currently, the following independent persons are in charge of the preparation of this proposal:

Sargon Hasso, R & D Programmer
Wallace Computer Services
10 Davis Dr.
Bellwood, IL 60104
(708) 547-8600 ext. 23

Peter Jasim, System Programmer
The Riverside Publishing Co.
8420 Bryn Mawr Ave
Chicago, IL 60631
(312) 369-0201

Edward Odisho, Ph.D.
Associate Professor
Northeastern Illinois University
5500 N. St. Louis Ave.
Chicago, IL 60625
(312) 794-2727

cc: PJ
EO

Assyrian¹ S²+0000 -- SSSS+0000

The Assyrian script (*Syriac*), which is used for writing the Assyrian language, includes the Eastern Assyrian script (*Nestorian*), the Western Assyrian script (*Serto* or *Jacobite*), and the *Estrangelo*³ script. It is also used for writing Mandaic⁴.

Assyrian script which is predominantly cursive is written from right to left even in its printed form. Few letters are written in different forms depending on how they join to their neighbors. Vowels (*zaw'e*) are placed above or below the consonantal base letters (radical).

Mandaic. Mandaic is written with the same script, with an additional, 23rd letter; this extra letter is given the independent code SSSS+0023.

Punctuation. Most punctuation marks used in Assyrian are not given independent codes (they are unified with the Latin, Arabic, and Hebrew punctuation) except for the few cases where the mark has a unique form and function in Assyrian.

Encoding Principles. The alphabet of Assyrian is well defined. Each letter receives only one Unicode character value regardless of the number of contextual shapes it may exhibit in text (this, indeed, is the only difference between Eastern Assyrian, Western Assyrian, Estrangela, and Mandaic). The graphic form (glyph) shown in the Unicode character chart is primarily that of free-form Estrangela.

*Diacritical Points*⁵ (*paroshe*). These are marks (more commonly, these are points of large, medium, and small sizes) that indicate vowels (*zaw'e*), cantillation marks, accents, and other modifications of consonantal letters. The occurrence of a character in the *Paroshe* range and its depiction in relation to a dashed circle constitute an assertion that this character is intended to be applied via some process to the consonantal letter, phrase and/or clause that precedes it in the text stream. General rules for applying non-spacing marks are given in the Generic Diacritical Mark block description section in the Unicode Standard, version 1.0. The Unicode standard does not specify a sequence order in case of multiple marks

¹ We use *Assyrian* and *Syriac* interchangeably in this working proposal. However, we would strongly suggest the use of *Assyrian* as a proper name for this script.

² This notation is for this working proposal only, and it follows the same convention as used in Unicode Version 1.0, i.e., U+nnnn. *S* stands for Syriac.

³ ~~The correct spelling of *Estrangelo* is with an *o* at the end, not an *a*.~~ It is also spelled with an *o* the end.

⁴ Two issues were raised in the exploratory proposal: the order of letters and Mandaic. The order of letters is correct as it appears -- *Waw* is in its correct place. *Waw* should not be placed at the end because *Waw* is, by virtue of its position in the Assyrian alphabet (the sixth letter) also the number 6. Assyrian letters are also used as numbers and have ordinal values. More information will be provided if need be. We have left the last space in the character set as reserved for the extra letter in Mandaic.

⁵ Generally we refer to all objects that are placed around the base letter in various positions, as *Diacritical Points*. However, they fall into four well-defined categories: distinction points, e.g., SSS+0005, the plural sign, e.g., SSS+0001, the actual diacritical point in its various forms e.g., SSS+0008, and, finally, the accents, e.g., SS+0008. For an in depth treatment of this subject, please cf. Segal's *The Diacritical Point and The Accents in Syriac*, Oxford University Press, 1953.

applied to the same Assyrian base character since there is no possible ambiguity of interpretation. The Assyrian script contains a rich set of diacritical marks which reflects its development over the course of its long history.













Encoding Structure. The Assyrian character block is divided into the following⁶:







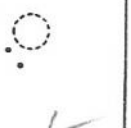


S+0001-- S+0010	Assyrian punctuation and number marks
SS+0001 -- SS+0020 ⁷	Assyrian diacritical marks (<i>paroshe</i>) -- I
SSS+0001 -- SSS+0017	Assyrian diacritical marks (<i>paroshe</i>) -- II
SSSS+0001 -- SSSS+0023	Assyrian letters















⁶ This subdivision is for future addition and expansion. Cf. following note.















⁷ This list by no means is complete. Although, the majority of all Assyrian documents and manuscripts can be reproduced with the aid of only this list -- whether they are from 500 A.D. or from 1993 A.D. Therefore, we would like to reserve additional space following this category for future addition and amendment.

SEQUENCE	GLYPH	DEFINITION
S+0001	¶	SYRIAC QUOTATION MARK = <i>sahrane</i>
S+0002	ⲥ	SYRIAC ABBREVIATION MARK = <i>gadmana</i>
S+0003	Ⲅ	SYRIAC END OF PARAGRAPH SEPARATOR
S+0004	.	SYRIAC LONG PAUSAL MARK PASUQA = <i>period</i>
S+0005	,	SYRIAC SHORT PAUSAL MARK ZAUGA = <i>comma</i>
S+0006	Ⲅ	SYRIAC QUESTION MARK
S+0007	Ⲁ	SYRIAC NUMERAL SIGN FOR TEN
S+0008	Ⲁ	SYRIAC NUMERAL SIGN FOR THOUSAND
S+0009	Ⲁ	SYRIAC NUMERAL SIGN FOR TEN THOUSAND
S+0010	Ⲁ	SYRIAC NUMERAL SIGN FOR MILLION
SS+0001	Ⲁ	SYRIAC ONE DOT ABOVE RIGHT = <i>mshalana</i>

SS+0002		SYRIAC TWO DOTS HORIZONTAL ABOVE = <i>rahta</i>
SS+0003		SYRIAC TWO DOTS LEFT-SLANTED ABOVE = <i>mshalmanota</i>
SS+0004		SYRIAC ONE DOT ABOVE CENTER = <i>ritma</i>
SS+0005		SYRIAC TWO DOTS VERTICAL CENTER ABOVE = <i>zauga elaya</i> 
SS+0006		SYRIAC THREE DOTS ABOVE = <i>rahta d karte</i>
SS+0007		SYRIAC ONE DOT ABOVE LEFT = <i>esyana</i>
SS+0008		SYRIAC TWO DOTS VERTICAL RIGHT ABOVE = <i>mtdamrana</i>
SS+0009		SYRIAC ONE DOT IN-LINE LEFT = <i>pasoga</i>
SS+0010		SYRIAC TWO DOTS VERTICAL IN-LINE LEFT = <i>zauga</i>
SS+0011		SYRIAC TWO DOTS IN-LINE ABOVE AND BELOW = <i>mqimana</i>
SS+0012		SYRIAC ONE DOT RIGHT BELOW = <i>mnakhta</i>

SS+0013		SYRIAC ONE DOT LEFT BELOW = <i>samka</i>
SS+0014		SYRIAC TWO DOTS VERTICAL LEFT BELOW = <i>mitkashpana</i>
SS+0015		SYRIAC ONE LARGE DOT AND ONE SMALL DOT SLANTED LEFT BELOW = <i>esasa</i>
SS+0016		SYRIAC ONE LARGE DOT AND ONE SMALL DOT BELOW = <i>napsha</i>
SS+0017		SYRIAC TWO DOTS RIGHT-SLANTED ABOVE LEFT = <i>elaya</i>
SS+0018		SYRIAC TWO DOTS ABOVE AND IN-LINE DOT LEFT = <i>rahta d paseq</i>
SS+0019		SYRIAC ONE SMALL DOT AND ONE LARGE DOT LEFT- SLANTED BELOW LEFT = <i>takhtaya</i>
SS+0020		SYRIAC THREE DOTS BELOW = <i>takhtaya d talata</i>
SSS+0001		SYRIAC PLURAL MARK SYAME

SSS+0002		SYRIAC FRICATION MARK RUKAKHA
SSS+0003		SYRIAC AFFRICATION MARK MAJLIANA
SSS+0004		SYRIAC HALF CIRCLE BELOW A LETTER = <i>qishta</i>
SSS+0005		SYRIAC DIACRITICAL LARGE POINT OVER A LETTER
SSS+0006		SYRIAC DIACRITICAL LARGE POINT BELOW A LETTER
SSS+0007		SYRIAC DIACRITICAL POINT QUSHAYA
SSS+0008		SYRIAC VOWEL MARK ZQAPA
SSS+0009		SYRIAC VOWEL MARK PTAKHA
SSS+0010		SYRIAC VOWEL MARK ZLAME PSHIQE
SSS+0011		SYRIAC VOWEL MARK ZLAME QASHYE
SSS+0012		SYRIAC VOWEL MARK RWAKHA
SSS+0013		SYRIAC VOWEL MARK RWASA
SSS+0014		SYRIAC VOWEL MARK KHWASA
SSS+0015		SYRIAC ACCENT MTALQANA

SSS+0016		SYRIAC HALF-VOWEL MHAGYANA
SSS+0017		SYRIAC HALF-VOWEL MARHTANA
SSSS+0001		SYRIAC LETTER ALLAP
SSSS+0002		SYRIAC LETTER BET
SSSS+0003		SYRIAC LETTER GAMMAL
SSSS+0004		SYRIAC LETTER DALLAT
SSSS+0005		SYRIAC LETTER HE
SSSS+0006		SYRIAC LETTER WAW
SSSS+0007		SYRIAC LETTER ZAIN
SSSS+0008		SYRIAC LETTER KHET
SSSS+0009		SYRIAC LETTER TET
SSSS+0010		SYRIAC LETTER YUDH
SSSS+0011		SYRIAC LETTER KAP
SSSS+0012		SYRIAC LETTER LAMMADH

SSSS+0013	𐎎	SYRIAC LETTER MEEM
SSSS+0014	𐎍	SYRIAC LETTER NUN
SSSS+0015	𐎌	SYRIAC LETTER SIMKAT
SSSS+0016	𐎋	SYRIAC LETTER 'E
SSSS+0017	𐎊	SYRIAC LETTER PE
SSSS+0018	𐎉	SYRIAC LETTER SADEH
SSSS+0019	𐎈	SYRIAC LETTER QOP
SSSS+0020	𐎇	SYRIAC LETTER RESH
SSSS+0021	𐎆	SYRIAC LETTER SHEEN
SSSS+0022	𐎅	SYRIAC LETTER TAW
SSSS+0023		RESERVED

Oct 19
November 4, 1993

Mr. Rick McGowan
NeXt, Inc.,
900 Chesapeake Drive
Redwood City, CA 94063

Dear Rick:

A while back I have spoken with you on the phone regarding the inclusion of the Syriac script in the Unicode standard. Shortly thereafter, Peter Jasim have had some correspondence with you regarding the same matter. You have indicated to me then that no one had submitted any proposal yet. Therefore, we are submitting this proposal to the Unicode Consortium for thorough examination and possible inclusion in the Unicode standard. Please feel free to contact me for any question and/or clarification that you, or the committee in charge, might have.

Sincerely,

Sargon Hasso, Independent Consultant

P.S.

Currently, the following independent persons are in charge of the preparation of this proposal:

Sargon Hasso, R & D Programmer
Wallace Computer Services
10 Davis Dr.
Bellwood, IL 60104
(708) 547-8600 ext. 23

Peter Jasim, System Programmer
The Riverside Publishing Co.
8420 Bryn Mawr Ave
Chicago, IL 60631
(312) ~~369-0201~~

Edward Odisho, Ph.D.
Associate Professor
Northeastern Illinois University
5500 N. St. Louis Ave.
Chicago, IL 60625
(312) 794-2727

cc: PJ
EO

Assyrian¹ S²+0000 -- SSSS+0000

The Assyrian script (*Syriac*) is used for writing the Assyrian language, which include the Eastern Assyrian script (*Nestorian*), the Western Assyrian script (*Serto* or *Jacobite*), and the *Estrangelo*³ script. It is also used for writing Mandaic⁴.

Assyrian is written from right to left, and is predominantly cursive, even in its printed form. Few letters are written in different forms, depending on how they join to their neighbors. Vowels are written as *zaw'e* applied to consonantal base letters.

Mandaic. Mandaic is written with the same script, with an additional, 23rd letter; this extra letter is given the independent code SSSS+0023.

Punctuation. Most punctuation marks used in Assyrian are not given independent codes (they are unified with the Latin, Arabic, and Hebrew punctuation), except for the few cases where the mark has a unique form and function in Assyrian.

Encoding Principles. The alphabet of Assyrian is well defined. Each letter receives only one Unicode character value, regardless of the number of contextual shapes it may exhibit in text (this, indeed, is the only difference between Eastern Assyrian, Western Assyrian, Estrangelo, and Mandaic). The graphic form (glyph) shown in the Unicode character chart, for the most part, is that of free-form Estrangelo.

*Diacritical Points*⁵ (*paroshe*). These are marks (more commonly, these are points of large, medium, and small sizes) that indicate vowels (*zaw'e*), cantillation marks, accents, and other modifications of consonantal letters. The occurrence of a character in the *Paroshe* range, and its depiction in relation to a dashed circle, constitute an assertion that this character is intended to be applied via some process to the consonantal letter, phrase and/or clause that precedes it in the text stream. General rules for applying non-spacing marks are given in the Generic Diacritical Mark block description section in the Unicode Standard, version 1.0. The Unicode standard does not specify a sequence order in case of multiple marks

¹ We use *Assyrian* and *Syriac* interchangeably in this working proposal. However, we would strongly suggest the use of *Assyrian* as a proper name for this script.

² This notation is for this working proposal only, and it follows the same convention as used in Unicode Version 1.0, i.e., U+nnnn. *S* stands for Syriac.

³ The correct spelling of *Estrangelo* is with an *o* at the end, not an *a*.

⁴ Two issues were raised in the exploratory proposal: the order of letters and Mandaic. The order of letters is correct as it appears -- *Wow* is in its correct place. *Wow* should not be placed at the end because *Wow* is, by virtue of its position in the Assyrian alphabet (the sixth letter) also the number 6. Assyrian letters are also used as numbers and have ordinal values. More information will be provided if need be. We have left the last space in the character set as reserved for the extra letter in Mandaic.

⁵ Generally we refer to all objects that are placed around the base letter in various positions, as *Diacritical Points*. However, they fall into four well-defined categories: distinction points, e.g., SSS+0005, the plural sign, e.g., SSS+0001, the actual diacritical point in its various forms e.g., SSS+0008, and, finally, the accents, e.g., SS+0008. For an in depth treatment of this subject, please cf. Segal's *The Diacritical Point and The Accents in Syriac*, Oxford University Press, 1953.

applied to the same Assyrian base character, since there is no possible ambiguity of interpretation. The Assyrian script contains a rich set of diacritical marks, which reflects its development over the course of its long history.















Encoding Structure. The Assyrian character block is divided into the following⁶:














S+0001-- S+0010	Assyrian punctuation and number marks
SS+0001 -- SS+0020 ⁷	Assyrian diacritical marks (<i>paroshe</i>) -- I
SSS+0001 -- SSS+0017	Assyrian diacritical marks (<i>paroshe</i>) -- II
SSSS+0001 -- SSSS+0023	Assyrian letters















⁶ This subdivision is for future addition and expansion. Cf. following note.

⁷ This list by no means is complete. Although, the majority of all Assyrian documents and manuscripts can be reproduced with the aid of only this list -- whether they are from 500 A.D. or from 1993 A.D. Therefore, we would like to reserve additional space following this category for future addition and amendment.

SEQUENCE	GLYPH	DEFINITION
S+0001	¶	SYRIAC QUOTATION MARK = <i>sahrane</i>
S+0002	‡	SYRIAC ABBREVIATION MARK = <i>gadmana</i>
S+0003	❖	SYRIAC END OF PARAGRAPH SEPARATOR
S+0004	.	SYRIAC LONG PAUSAL MARK PASUQA = <i>period</i>
S+0005	:	SYRIAC SHORT PAUSAL MARK ZAUGA = <i>comma</i>
S+0006	؛	SYRIAC QUESTION MARK
S+0007	⦿	SYRIAC NUMERAL SIGN FOR TEN
S+0008	⦿/⦿	SYRIAC NUMERAL SIGN FOR THOUSAND
S+0009	⦿ 	SYRIAC NUMERAL SIGN FOR TEN THOUSAND
S+0010	⦿ ^	SYRIAC NUMERAL SIGN FOR MILLION
SS+0001	⦿ [•]	SYRIAC ONE DOT ABOVE RIGHT = <i>mshalana</i>

SS+0002		SYRIAC TWO DOTS HORIZONTAL ABOVE = rahta
SS+0003		SYRIAC TWO DOTS LEFT-SLANTED ABOVE = mshalmanota
SS+0004		SYRIAC ONE DOT ABOVE CENTER = ritma
SS+0005		SYRIAC TWO DOTS VERTICAL CENTER ABOVE = zauga elaya
SS+0006		SYRIAC THREE DOTS ABOVE = rahta d karte
SS+0007		SYRIAC ONE DOT ABOVE LEFT = esyana
SS+0008		SYRIAC TWO DOTS VERTICAL RIGHT ABOVE = mtdamrana
SS+0009		SYRIAC ONE DOT IN-LINE LEFT = pasoqa
SS+0010		SYRIAC TWO DOTS VERTICAL IN-LINE LEFT = zauga
SS+0011		SYRIAC TWO DOTS IN-LINE ABOVE AND BELOW = mqimana
SS+0012		SYRIAC ONE DOT RIGHT BELOW = mnakhta
SS+0013		SYRIAC ONE DOT LEFT BELOW = samka
SS+0014		SYRIAC TWO DOTS VERTICAL LEFT BELOW = mitkashpana
SS+0015		SYRIAC ONE LARGE DOT AND ONE SMALL DOT SLANTED LEFT BELOW = esasa

SS+0016		SYRIAC ONE LARGE DOT AND ONE SMALL DOT BELOW = napsha
SS+0017		SYRIAC TWO DOTS RIGHT-SLANTED ABOVE LEFT = elaya
SS+0018		SYRIAC TWO DOTS ABOVE AND IN-LINE DOT LEFT = rahta d paseq
SS+0019		SYRIAC ONE SMALL DOT AND ONE LARGE DOT LEFT-SLANTED BELOW LEFT = takhtaya
SS+0020		SYRIAC THREE DOTS BELOW = takhtaya d talata
SSS+0001		SYRIAC PLURAL MARK SYAME
SSS+0002		SYRIAC FRICATION MARK RUKAKHA
SSS+0003		SYRIAC AFFRICATION MARK MAJLIANA
SSS+0004		SYRIAC HALF CIRCLE BELOW A LETTER = qishta
SSS+0005		SYRIAC DIACRITICAL LARGE POINT OVER A LETTER
SSS+0006		SYRIAC DIACRITICAL LARGE POINT BELOW A LETTER
SSS+0007		SYRIAC DIACRITICAL POINT QUSHAYA
SSS+0008		SYRIAC VOWEL MARK ZQAPA

SSS+0009		SYRIAC VOWEL MARK PTAKHA
SSS+0010		SYRIAC VOWEL MARK ZLAME PSHIQE
SSS+0011		SYRIAC VOWEL MARK ZLAME QASHYE
SSS+0012		SYRIAC VOWEL MARK RWAKHA
SSS+0013		SYRIAC VOWEL MARK RWASA
SSS+0014		SYRIAC VOWEL MARK KHWASA
SSS+0015		SYRIAC ACCENT MTALQANA
SSS+0016		SYRIAC HALF-VOWEL MHAGYANA
SSS+0017		SYRIAC HALF-VOWEL MARHTANA
SSSS+0001		SYRIAC LETTER ALLAP
SSSS+0002		SYRIAC LETTER BET
SSSS+0003		SYRIAC LETTER GAMMAL
SSSS+0004		SYRIAC LETTER DALLAT
SSSS+0005		SYRIAC LETTER HE

SSSS+0006	ܘ	SYRIAC LETTER WAW
SSSS+0007	ܙ	SYRIAC LETTER ZAIN
SSSS+0008	ܠ	SYRIAC LETTER KHET
SSSS+0009	ܡ	SYRIAC LETTER TET
SSSS+0010	ܢ	SYRIAC LETTER YUDH
SSSS+0011	ܟ	SYRIAC LETTER KAP
SSSS+0012	ܠܡܡܘܕܗ	SYRIAC LETTER LAMMADH
SSSS+0013	ܡܡ	SYRIAC LETTER MEEM
SSSS+0014	ܢܢ	SYRIAC LETTER NUN
SSSS+0015	ܫܫܟܬܐ	SYRIAC LETTER SIMKAT
SSSS+0016	ܐܐ	SYRIAC LETTER E
SSSS+0017	ܦܦ	SYRIAC LETTER PE
SSSS+0018	ܫܫܘܕܗ	SYRIAC LETTER SADEH
SSSS+0019	ܩܩ	SYRIAC LETTER QOP
SSSS+0020	ܠܫܐ	SYRIAC LETTER RESH
SSSS+0021	ܫܫܝܢ	SYRIAC LETTER SHEEN
SSSS+0022	ܬܬܐ	SYRIAC LETTER TAW

SSSS+0023		RESERVED
-----------	--	----------

10/11/93



Next, Inc.
900 Chesapeake Drive
Redwood City, CA 94063

Peter Jasim
Nineveh Software
2317 West Farwell
Chicago, IL 60645



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	ܚ	ܝ					ܘ	ܝ
4	ܒ	ܚ					ܘ	ܝ
5	ܘ	ܝ					ܘ	ܝ
6	ܘ						ܘ	ܝ
7	ܚ						ܘ	
8	ܚ						ܘ	
9	ܘ						ܘ	
A	ܘ						ܘ	
B	ܘ						ܘ	
C	ܚ						ܘ	
D	ܘ						ܘ	
E	ܘ						ܘ	
F	ܘ						ܘ	

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



George A. Kiraz
St. John's College
Cambridge CB2 1TP
United Kingdom

Dear George:

The last time I wrote to it was in regard to the Ashurbanipal Library Computer Conference. This time I am writing to you in regard to The First Forum on Syriac Computing. How small is the world of Syriac and computers.

Perhaps you may remember Sargon Hasso, whom you met at the Ashurbanipal Library Computer Conference. He and I co-authored the Syriac Unicode standard, and have submitted this standard to the Unicode Consortium. Unicode will in turn submit this standard to the International Standards Organization (ISO) for ratification. I wish to present a paper at SyrCOM about this important development. Unicode has become the de facto standard for the next generation of computers and computing. It is supported by a wide variety of operating systems and computer vendors. When the Syriac Unicode standard is approved, Syriac scholars will be able to electronically exchange documents, scientific papers, and electronic forms of manuscripts (which opens a vast, new realm of electronic processing of Syriac text) with complete compatibility.

Enclosed you will find an abstract for the paper I wish to present. Please give it your utmost consideration.

Yours in Assyrianism

Peter Jasim



The Syriac Unicode Standard
















by

Peter Jasim
Sargon Hasso

abstract

The American Standard Code for Information Interchange (ASCII), a standard encoding of the letters, digits, and symbols of the English language, will soon be replaced by the Unicode Standard. ASCII is a one byte coding scheme, which provides enough space for 256 codes, hence it cannot encode languages with more than 256 letters, digits, and symbols. Various multibyte encoding schemes have been proposed, but it is the Unicode Standard that has emerged, principally because of its simplicity, orthogonality, and its comprehensive scope. Unicode encodes all of the major and minor languages of the world, including Syriac, in a two byte character set. This paper presents the Syriac Unicode Standard which has been submitted to the Unicode Consortium, and which is awaiting ratification by the International Standards Organization. A brief overview of the philosophy and design of the Unicode Standard is presented, the unification process which lead to the Syriac Standard is discussed, and the Syriac Standard is presented. The Syriac Standard unifies the Western ("Jacobite") and Eastern ("Nestorian") character sets, and also supports Mandaic. With the Syriac Unicode Standard, any document produced between 0 A.D. and 1995 A.D. can be represented.

SEQUENCE	GLYPH	DEFINITION
S+0000	¶	SYRIAC QUOTATION SIGN = <i>sahrane</i>
S+0001	⋈	SYRIAC ABBREVIATOR SIGN
S+0002	⋄	SYRIAC END OF PARAGRAPH SIGN
S+0003	.	SYRIAC PAUSAL ACCENT PASUQA = <i>period</i>
S+0004	,	SYRIAC PAUSAL ACCENT ZAUGA = <i>comma</i>
S+0005	؟	SYRIAC INTERROGATIVE SIGN
S+0006	⋈	SYRIAC PLURAL SIGN SYAME
S+0007	⋈	SYRIAC DIACRITICAL POINT RUKAKHA
S+0008	⋈	SYRIAC DIACRITICAL SIGN MAJLIANA
S+0009	⋈	SYRIAC DIACRITICAL LARGE POINT OVER A LETTER
S+000A	⋈	SYRIAC DIACRITICAL LARGE POINT BELOW A LETTER
S+000B	⋈	SYRIAC DIACRITICAL POINT QUSHAYA
S+000C	⋈	SYRIAC VOWEL SIGN ZQAPA
S+000D	⋈	SYRIAC VOWEL SIGN PTAKHA
S+000E	⋈	SYRIAC VOWEL SIGN ZLAME PSHIQE

S+000F		SYRIAC VOWEL SIGN ZLAME QASHYE
S+0010		SYRIAC VOWEL SIGN RWAKHA
S+0011		SYRIAC VOWEL SIGN RWASA
S+0012		SYRIAC VOWEL SIGN KHWASA
S+0013		SYRIAC ACCENT MASHLIANA
S+0014		SYRIAC ACCENT MHAGYANA
S+0015		SYRIAC ACCENT MARHTANA
S+0016		SYRIAC LETTER ALAP
S+0017		SYRIAC LETTER BET
S+0018		SYRIAC LETTER GAMAL
S+0019		SYRIAC LETTER DALADH
S+001A		SYRIAC LETTER HE
S+001B		SYRIAC LETTER WAW
S+001C		SYRIAC LETTER ZAIN
S+001D		SYRIAC LETTER HET

S+001E	ܦ	SYRIAC LETTER TET
S+001F	ܦ	SYRIAC LETTER YUDH
S+0020	ܦ	SYRIAC LETTER KAP
S+0021	ܦ	SYRIAC LETTER LAMADH
S+0022	ܦ	SYRIAC LETTER MIM
S+0023	ܦ	SYRIAC LETTER NUN
S+0024	ܦ	SYRIAC LETTER SEMKAT
S+0025	ܦ	SYRIAC LETTER 'E
S+0026	ܦ	SYRIAC LETTER PE
S+0027	ܦ	SYRIAC LETTER SADE
S+0028	ܦ	SYRIAC LETTER QOP
S+0029	ܦ	SYRIAC LETTER RESH
S+002A	ܦ	SYRIAC LETTER SHEEN
S+002B	ܦ	SYRIAC LETTER TAW

Seventh To Tenth Centuries, Eastern

- | | | |
|------|---|--|
| ✓ 1 | ◦ | 2Aaeabab, 2ayt, 2abab 2abab |
| 2 | ◦ | 2abab, 2ayt |
| ✓ 3 | ◦ | 2abab, 2ayt, 2abab |
| ✓ 4 | ◦ | 2abab |
| ✓ 5 | ◦ | 2Aaeabab, 2ayt |
| ✓ 6 | ◦ | 2ayt, 2abab, 2abab, 2ayt, 2abab, 2ayt |
| ✓ 7 | ◦ | 2abab |
| ✓ 8 | ◦ | 2ayt |
| | ◦ | 2ayt |
| | ◦ | 2ayt |
| ✓ 9 | ◦ | 2ayt, 2abab |
| ✓ 10 | ◦ | 2ayt, 2abab |
| ✓ 11 | ◦ | 2abab |
| ✓ 12 | ◦ | 2ayt, 2abab, 2abab |
| 13 | ◦ | 2ayt |
| ✓ 14 | ◦ | 2ayt |
| ✓ 15 | ◦ | 2abab |
| ✓ 16 | ◦ | 2abab |
| ✓ 17 | ◦ | 2abab |
| ✓ 18 | ◦ | 2abab, 2abab |
| ✓ 19 | ◦ | 2abab, 2abab |
| ✓ 20 | ◦ | 2ayt |

EASTERN

21

:0

247016

22

0
:

247016

Seventh to Tenth Centuries, Western

- 1 0 2000
- 2 0 2000, 2100
- ✓3 0 2000, 2100, 2200, 2300
 2400, 2500, 2600
- ✓4 :0 2100
- ✓5 0 2000
- ✓6 0 2000
- ✓7 0 | 0 2000, 2100

- ✓8 0 2000, 2100
- ✓9 0 2000, 2100, 2200
 2300, 2400

- ✓10 :0 2000, 2100, 2200
- ✓11 0 2000, 2100
- ✓12 :0 2000, 2100

Before
Seventh Century A 64-65

- طحلا لعل، كصم و ع
- لرمع
- وقتك للنت (طوسل) ، كصم (طاذجلا)
- كصم

- طاذجلا ، طوسل
- كصم ، طاذجلا

- كصم
- لعل
- كصم
- وقتك (طوسل) ، كصم (طاذجلا)

- طاذجلا

UNIFIED DIACRITICAL MARKS

					EASTERN	WESTERN
U	1	Ö ✓		2a12ab	3	3
U	2	ö ✓		2ayt	1	1/2
U	3	ö o ✓		2aa1 2a12ab 2aog	6	5/6
U	4	ö ö ö ✓		2a12ab	8	7
B	5	o ✓		2aueb 2a12ab	12	9
B	6	o ✓		2abab	11 (13)	8
L	7	.o ✓		2aaad	19	11
UL	7	.ö ✓	7+2 (?)	2ll	7	4
BL	9	.o ✓	7+6 (?)	2a12ab	17	10
L	10	:o ✓		2og	20	12
B	11	o ✓		2a12ab	22	
U	12	ö ✓		2aa12	2	
U	13	o: ✓		2a12ab	4	
U	14	ö ✓		2aaab12ab 2ayt	5	
U	15	ö ✓		2aa12ab 2a12ab	9	
UL	16	.ö ✓	(4+7)?	2aa12ab 2a12ab	10	
B	17	o ✓		2ayt	14	
B	18	o o ✓		2aa1	15/16	
LB	19	:o	7+9	2aa12ab 2aa12ab	18	
L	20	:o ✓		2abab	21	