

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

Doc Type: Working Group Document
Title: EGYPTOLOGICAL YOD and Cyrillic breathings
Source: Michael Everson and Bob Richmond
Status: Individual Contribution
Date: 2008-01-28
Replaces: N2241 (2000-08-27)

Request. This document asks for the change of the property of U+0485 COMBINING CYRILLIC DASIA PNEUMATA and U+0486 COMBINING CYRILLIC PSILI PNEUMATA from “Script=Cyrillic” to “Script=Inherited”. It also asks for the addition of the following informative notes to those two characters.

- ◌̇ 0485 COMBINING CYRILLIC DASIA PNEUMATA
- rests atop lowercase and uppercase Cyrillic and Greek letters
 - rests atop lowercase Latin letter but precedes uppercase Latin letters
 - in origin the spiritus asper
- x 0314 combining reversed comma above
- ◌̈ 0486 COMBINING CYRILLIC PSILI PNEUMATA
- rests atop lowercase and uppercase Cyrillic and Greek letters
 - rests atop lowercase Latin letter but precedes uppercase Latin letters
 - used with i for Egyptological yod
 - used with a and i and u in transliteration of Ugaritic
 - in origin the spiritus lenis
- x 0313 combining comma above

Rationale. The existing U+0314 COMBINING REVERSED COMMA ABOVE and U+0313 COMBINING COMMA ABOVE, when used with Latin, Greek, and Cyrillic scripts, have a specifically-defined glyph representation behaviour. These two characters cannot be used with the Latin script for Egyptological and Ugariticist use because the required glyph representation behaviour is different. The two Cyrillic combining characters here, however, can be used to solve this problem rather neatly. The examples given in the Latin column are the three vowels used in Egyptology and Ugaritic studies as well as letters used in a North Wakashan language of the Pacific Northwest. Some minority languages which use the Cyrillic script do, or did, use U+0313 centred above as shown. The behaviour of Greek for U+0486 is given for symmetry, though we do not insist that its behaviour with U+0486 be specified.

	Latin	Greek	Cyrillic
U+0313	À Í Û Ñ Ñ Ñ Ñ	Α Ι Ύ Ρ	À Ñ Ù Ñ Ñ Ñ
U+0313	à í û ñ ñ ñ ñ	α ι υ ρ	à ñ ù ñ ñ ñ
U+0486	À Í Û Ñ Ñ Ñ Ñ	À Í Ύ Ρ	À Ñ Ù Ñ Ñ Ñ
U+0486	à í û ñ ñ ñ ñ	α ι υ ρ	à ñ ù ñ ñ ñ

Implementation. COMBINING CYRILLIC PSILI PNEUMATA can be used already with some software, most notably Microsoft Office (as tested with versions 2003 and 2007). However “Script=Cyrillic” is often interpreted as “only apply the modifier to a Cyrillic base character”. Tests on Firefox, Internet Explorer, and Safari revealed this behaviour, in that the modifier is not applied even when a suitable (TrueType) font is specified. WordPerfect and OpenOffice likewise do not allow the modifier to take effect. A change to “Script=Inherited” would state that the modifier should be applied to Latin characters.

Bibliography

- Gardiner, Alan. 1966. *Egyptian grammar: being an introduction to the study of hieroglyphs*. 3rd edition. London: Oxford University Press.
- Hetzron, Robert, ed. 1997. *The Semitic Languages*. London: Routledge.
- Ritter, R. M. 2002. *The Oxford guide to style*. Oxford: Oxford University Press. ISBN 0-19-869175-0

Examples.

𐎛 *ini* (W 25) bring, fetch, remove; 𐎛𐎠𐎠𐎠
 var. 𐎛𐎠𐎠 *inw* gifts, tribute; 𐎛𐎠𐎠𐎠 varr.
 𐎛𐎠𐎠, 𐎛𐎠 *In-hrt* (N 31) Onūris, the god
 of This, N. of Abydus, Gk. *’Ovoūpis*.

Figure 1. Example from Gardiner 1966, showing *i* with *spiritus lenis* in *In-hrt* and LATIN SMALL LETTER LENIS I in *ini* and *inw*.

Special characters are *hamza* ʾ (capital) and ʾ (lower case) (sometimes approximated in typescript by a 3), and *ʿayn* ʿ (capital) and ʿ (lower case) (sometimes approximated in typescript by a 9), *ʾI i* (*alef lenis* combined with *I* and *i*, often called the ‘*lenis-i*’). The Ancient Egyptian *alef* and *ʿayn* are both larger than the corresponding Arabic characters, and have capital and lower-case forms: ʿ ʿ ʾ ʾ *I i*.

Figure 2. Example from Ritter 2002, showing three characters used in Egyptology with casing pairs.

The Abecedaries and the Consonantal Alphabet

The order of the alphabet is known from abecedaries and is similar to that of the later Northwest Semitic languages:

UG	ʾ	b	g	ḥ	d	h	w	z	ḥ	ṭ	y	k	š	l	m	ḏ	n	z	s	ʿ	p	š	q	r	θ	γ	t	ʾ	ù	š
NWS	ʾ	b	g		d	h	w	z	ḥ	ṭ	y	k		l	m		n		s	ʿ	p	š	q	r	š		t			

The basic consonantal inventory consisted of twenty-seven phonemes; the origin of the last three signs is in dispute. The three *ʾaleph* signs are used to indicate /ʾ/ plus following vowel (e.g., ⟨ʾà⟩ = /ʾa/), with ⟨ʾi⟩ used for syllable-final /ʾ/.

Figure 3. Chart from Hetzron 1997 showing Ugaritic vowels with *spiritus lenis*.

Table 37. Northwest Semitic Scripts

Value	Ugaritic	Hebrew	Estrangelo	Serto	Nestorian	Numeric Value
ʾ (á)	𐎀	א	ܐ	Ⲁ	Ⲁ	1
b	𐎁	ב	ܒ	Ⲃ	Ⲃ	2
g	𐎂	ג	ܓ	Ⲅ	Ⲅ	3
(h)	𐎃					
d	𐎄	ד	ܕ	Ⲇ	Ⲇ	4
h	𐎅	ה	ܗ	Ⲉ	Ⲉ	5
w	𐎆	ו	ܘ	Ⲋ	Ⲋ	6
z	𐎇	ז	ܙ	Ⲍ	Ⲍ	7
h	𐎈	ח	ܚ	Ⲏ	Ⲏ	8
t	𐎉	ט	ܛ	Ⲑ	Ⲑ	9
y	𐎊	י	ܝ	Ⲓ	Ⲓ	10
k	𐎋	כ	ܟ	Ⲕ	Ⲕ	20
(š)	𐎌					
l	𐎍	ל	ܠ	Ⲗ	Ⲗ	30
m	𐎎	מ	ܡ	Ⲙ	Ⲙ	40
(ð)	𐎏					
n	𐎐	נ	ܢ	Ⲛ	Ⲛ	50
z	𐎑					
s	𐎒	ס	ܣ	Ⲝ	Ⲝ	60
c	𐎓	ס	ܥ	Ⲟ	Ⲟ	70
p	𐎔	פ	ܦ	Ⲙ	Ⲙ	80
š	𐎕	ש	ܨ	Ⲛ	Ⲛ	90
q	𐎖	ק	ܩ	Ⲝ	Ⲝ	100
r	𐎗	ר	ܪ	Ⲟ	Ⲟ	200
š	𐎘	ש	ܫ	Ⲙ	Ⲙ	300
š (θ)	𐎙	ש	ܫ	Ⲙ	Ⲙ	300
(γ)	𐎚					
t	𐎛	ת	ܬ	Ⲛ	Ⲛ	400
(i)	𐎜					
(ú)	𐎝					
(š)	𐎞					

Where two forms are shown, the one on the right occurs at the end of a word. Ugaritic values are given in parentheses.

Figure 4. Chart showing Ugaritic vowels with *spiritus lenis*.