URW Palladio HOT

Hybrid OpenType Diacritics Font for ISO 15919

Version 2.0 - 10th March 2004

http://www.sanskritweb.net

U.Stiehl@t-online.de

URW Palladio HOT
New Features of Version 2.0

The <u>original</u> version of the font URW Palladio HOT was completed on <u>January 1st, 2003</u>. The <u>new version</u> of this font completed on <u>March 10th, 2004</u> features these improvements:

1. Manual Hints

Great efforts were made to improve the screen legibility of the font by manual hinting, although hinting of Renaissance typefaces such as URW Palladio HOT is highly intricate.

2. Small Capitals

To the 4 styles (Roman, Bold, Italic, and Bold Italic) was added as 5th style for bookwork: Roman Small Capitals. This style is useful for emphasizing names in bibliographies etc. Germans will appreciate that "Gauß" converts to Gauss, if marked with Palladio HOT SC.

Note that "URW Palladio HOT SC" is a <u>separate font</u>. If I had included SC OpenType tables into the Roman style, Windows programs would not be able to access the small caps.

3. Additional Metrics Symbols

At the request of a German indologist, the following metrics/prosody signs were added:

Hexadecimal	F1F8	F1F9	F1FA	F1FB	F1FC	F1FD	F1FE	F1FF
Decimal	61944	61945	61946	61947	61948	61949	61950	61951
Metrics Symbol	˘	–	ᴜ̆	ꭒ	᷄	᷉	x́	x̀

These eight symbols are available as (non-italic) <u>regular cut</u> and (non-italic) <u>bold cut</u> only. Italic makes no sense here. If – inadvertantly – italic or bold italic is selected, the non-italic regular and bold versions of the metrical symbols will be displayed automatically. All 8 symbols have the same mono-spaced set width in both regular and bold cuts to permit of tabular composition without setting tabs, i.e. no tabs are required.

<u>Example:</u> āsīd rājā | nălō nāmă | vīrăsēnă-|sŭtō bălī: ‒‒‒‒|◡‒‒◡|‒◡‒◡|◡‒◡‒

4. Additional Chinese Diacritics

At the request of a researcher in East Asian Buddhism, 16 <u>additional</u> diacritics (= Unicode range 01CD–01DC hex) were added to allow for proper Pinyin transliteration of Chinese: Pinyin requires a total of 24 diacritical tone marks:

Tone 1	Ā, ā	Ē, ē	Ī, ī	Ō, ō	Ū, ū	Ǖ, ǖ
Tone 2	Á, á	É, é	Í, í	Ó, ó	Ú, ú	Ǘ, ǘ
Tone 3	Ǎ, ǎ	Ě, ě	Ǐ, ǐ	Ǒ, ǒ	Ǔ, ǔ	Ǚ, ǚ
Tone 4	À, à	È, è	Ì, ì	Ò, ò	Ù, ù	Ǜ, ǜ

<u>List of the 16 additional Pinyin diacritics in Unicode range 01CD–01DC</u>

Diacritic	Unicode hex.	Unicode decimal	Stacking Order
Ǎ	01CD	461	Aˇ
ǎ	01CE	462	aˇ
Ǐ	01CF	463	Iˇ
ǐ	01D0	464	iˇ
Ǒ	01D1	465	Oˇ
ǒ	01D2	466	oˇ
Ǔ	01D3	467	Uˇ
ǔ	01D4	468	uˇ
Ǖ	01D5	469	U¨¯
ǖ	01D6	470	u¨¯
Ǘ	01D7	471	U¨´
ǘ	01D8	472	u¨´
Ǚ	01D9	473	U¨ˇ
ǚ	01DA	474	u¨ˇ
Ǜ	01DB	475	U¨`
ǜ	01DC	476	u¨`

<u>Example:</u> "dào kě dào fēi cháng dào" (beginning of "Dàodé jīng" = "Tao Te King").

<u>Note:</u> Capital letters with two diacritics (Ǖ, Ǘ, Ǚ, Ǜ) require <u>leading above the lines</u> of text for proper (i.e. untruncated) display of the two stacked tone marks.

5. Macintosh Pies

To enable lossless exchange of files between Macintosh and Windows PC, seven missing pie characters were added to complete the Macintosh Code Pages: ∞ ∏ π ∫ Ω ≈

The new version 2.0 of URW Palladio HOT now covers a total of six Macintosh Code Pages: "US Roman", "Central Europe", "Croatian", "Icelandic", "Romanian", and "Turkish".

6. PUA Accents

The 27 Private Use Area accents in range F1D0-F1EA (in the "URW Palladio HOT Manual" always marked as "NEVER use these accents! Required internally only!") were removed to keep the font file size as compact as possible despite the enlargements made.

7. Additional Diacritics

The nasal diacritics A͠I, a͠ı (e.g. "ma͠ı" = "me") and A͞U, a͞u (e.g. "ca͞uknā" = "to be astonished") used in the Hindi-German dictionary by M. Gatzlaff-Hälsig were added.

Furthermore, the diacritics Ṙ,ṙ used in transliteration of Bengali were added.

The following URW Palladio HOT Manual was rewritten to reflect all above modifications.

<div style="border:1px solid red; color:red; text-align:center">

Note: The font URW Palladio HOT to which I have devoted so many months of hard work is not available to the general public and must not be offered for download on the internet. It is only available to my friends and colleagues.

</div>

Heidelberg, 10th March 2004   Ulrich Stiehl

URW Palladio HOT Manual

Hybrid OpenType Diacritics Font for ISO 15919

URW Palladio HOT is an innovative hybrid font (= two-in-one font) that can be used either as a plain Unicode font or as an OpenType Unicode font with glyph substitution features.

The character set of URW Palladio HOT, comprising 572 glyphs, covers the following:

1. All diacritics of the 4 Windows Code Pages for Pan-European Latin alphabets – that is: 1252 (Western Europe), 1250 (Central and Eastern Europe), 1257 (Baltic), 1254 (Turkish) – and many other diacritics of modern languages using the Latin alphabet.

   The new version 2004 of the font includes all diacritics of all Macintosh Code Pages too.

2. All diacritics required for ISO 15919 ("Transliteration of Devanāgarī and related Indic scripts into Latin characters") and many other non-ISO 15919 diacritics used in indology.

1. ISO 15919 Diacritics

The following chart shows the most often used diacritics for the transliteration of Sanskrit:

अ	a		आ	ā		इ	i		ई	ī		ळ	ḷ *	ḷ
उ	u		ऊ	ū		ऋ	r̥ *	r̥	ॠ	r̥̄ *	r̥̄	ॡ	ḹ *	ḹ
ए	e		ऐ	ai		ओ	o		औ	au		ऴ	ḷ	ḷ
क्	k		ख्	kh	kʰ	ग्	g		घ्	gh	gʰ	ङ्	ṅ	
च्	c		छ्	ch	cʰ	ज्	j		झ्	jh	jʰ	ञ्	ñ	
ट्	ṭ		ठ्	ṭh	ṭʰ	ड्	ḍ		ढ्	ḍh	ḍʰ	ण्	ṇ	
त्	t		थ्	th	tʰ	द्	d		ध्	dh	dʰ	न्	n	
प्	p		फ्	ph	pʰ	ब्	b		भ्	bh	bʰ	म्	m	
य्	y		र्	r		ल्	l		व्	v				
श्	ś		ष्	ṣ		स्	s		white = ISO 15919 diacritics					
ह्	h								gray = non-ISO 15919 diacritics					
ः	ḥ		˙	ṁ	ṃ	˚	m̐ *	m̐	* = not defined by Unicode (PUA)					

The complete ISO 15919 character set comprises the following alphabetical characters:

Aa Āā Ăă Àà Áá Ãã Ǎǎ Áá Ãã Ææ Ǣǣ Bb Cc Ĉĉ Dd Ḍḍ Ee Ēē Èè Éé Ẽẽ Êê Ềề Éé Ễễ Ff Gg Ġġ Hh H̲h̲ Ḥḥ Ḥ̱ḥ̱ Ḥ̄ḥ̄ Ii Īī Ìì Íí Ĩĩ Ìì Íí Ĩĩ Jj Kk K̲k̲ K̲h̲k̲h̲K̲H̲ Ll Ḷḷ Ĺḷ Ḷḷ Ḹḷ Ḹḷ Ìḷ Ĺḷ Mm Ṁṁ Ṃṃ M̌m̌ M̆m̆ Nn N̲n̲ Ňň Ññ Ṅṅ Ṇṇ Oo Ōō Òò Óó Õõ Ôô Ồồ Ốố Ỗỗ Pp Qq Rr R̲r̲ Řř Ṛṛ R̥r̥ R̥̄r̥̄ R̥̀r̥̀ Ŕ̥ŕ̥ R̥̀r̥̀ Ŕ̥ŕ̥ Ss Śś Ṣṣ Ṣ̱ṣ̱ S̲s̲ Tt Ṯṯ Ṭṭ Ṭ̱ṭ̱ Uu Ūū Ŭŭ Ùù Úú Ũũ Ùù Úú Ũũ Vv Ww Yy Ẏẏ Zz Z̲z̲ Żż Ẕẕ Žž

Each character of the ISO 15919 <u>alphabetical</u> character set is available in both lowercase and uppercase. The ligatures <u>Kh</u> <u>kh</u> <u>KH</u> are available in three versions: <u>Kh</u> (at the beginning of a sentence), <u>kh</u> (in the middle of a sentence), and <u>KH</u> (for capitalized chapter headings).

The <u>non-alphabetical</u> characters of ISO 15919 comprise figures (0–9), spaces and avagraha, usually encoded as apostrophe (0027), but in ISO 15919 encoded as single left quote (2018).

a) For <u>Classical Sanskrit</u>, at least the following ISO 15919 characters must be used:

  Aa Āā Bb Cc Dd Ḍḍ Ee Gg Hh Ḥḥ Ii Īī Jj Kk Ll Ḻḻ Mm Ṁṁ Nn Ññ Ṅṅ Ṇṇ Oo Pp Rr R̥r̥ R̥̄r̥̄ Ss Śś Ṣṣ Tt Ṭṭ Uu Ūū Vv Yy

b) For <u>Accented Vedic Sanskrit</u>, the following <u>additional</u> ISO 15919 characters may be used:

  Àà Áá Ầ ầ Ấấ Èè Éé Ìì Íí Ǐ̀ì̀ Íí Òò Óó R̥̀r̥̀ Ŕ̥ŕ̥ R̥̄́r̥̄́ Ùù Úú Ừừ Úú, and Ḻ̀ḻ̀ Ḻ́ḻ́ R̥̀̄r̥̀̄ (attestable?)

c) For <u>Classical/Vedic Sanskrit</u>, the following <u>additional</u> ISO 15919 characters may be used:

  Ḫḫ Ḫ̣ḫ̣ (old visargas), M̌m̌ (anunāsika, e.g. tām̌l lokān), Ḷḷ (intervocalic Ḍḍ), Ḹḹ (attested?)

The remaining ISO 15919 characters, i.e.

  Ăă Ãã Ẫẫ Ææ Ǣǣ Ĉĉ Ēē Ẽẽ Êê Èè Éé Ễễ Ff Ġġ Ḫḫ Ĭĭ Ĩĩ <u>Kk</u> <u>KhkhKH</u> Ḷḷ Ṃṃ M̌m̌ <u>Nn</u> Ňň Ōō Õõ Ôô Ồồ Ốố Ỗỗ Qq <u>Rr</u> Řř Ṛṛ Ṣṣ <u>Ss</u> Tt Ṭṭ Ŭŭ Ũũ Ữữ Ww Ẏẏ Zz <u>Zz</u> Żż Ẓẓ Žž

are used for other Indic languages (e.g. Marathi) and for other Indic scripts (e.g. Tamil), e.g.

  Sinhala: *Ææ Ǣǣ Ňň M̌m̌* etc.

  Arabic: Ff Ġġ Ḫḫ <u>KhkhKH</u> Qq Ṣṣ <u>Ss</u> Ṭṭ Ww Zz Żż Žž etc.

For a description of all the other diacritics (required for languages other than Sanskrit) visit Anthony Stone's homepage: http://homepage.ntlworld.com/stone-catend/translit.htm

<u>Caveats for Sanskritists</u>

1. Digraph characters, defined by ISO 15919, are <u>no</u> ligatures, but rather two simple letters: AI, ai; AU, au; KH, kh; GH, gh; CH, ch; JH, jh; ŇJ, ňj; ṬH, ṭh; ḌH, ḍh; ṚH, ṛh; ŇḌ, ňḍ; TH, th; DH, dh; ŇD, ňd; PH, ph; BH, bh; M̌B, m̌b. Note: ISO 15919 does <u>not</u> use kʰ etc.

2. "Macron below" for anudatta is <u>not</u> prescribed by ISO 15919. Therefore these diacritics have not been designed for URW Palladio HOT. Should you need these, use underline: <u>A</u>, <u>a</u>; <u>Ā</u>, <u>ā</u>; <u>I</u>, <u>i</u>; <u>Ī</u>, <u>ī</u>; <u>U</u>, <u>u</u>; <u>Ū</u>, <u>ū</u>; <u>R̥</u>, <u>r̥</u>; <u>R̥̄</u>, <u>r̥̄</u>; <u>Ḷ</u>, <u>ḷ</u>; <u>Ḹ</u>, <u>ḹ</u>; <u>E</u>, <u>e</u>; <u>Ē</u>, <u>ē</u>; <u>O</u>, <u>o</u>; <u>Ō</u>, <u>ō</u>

3. The "dot-below" diacritics ṛ, ḷ, ṃ are used by ISO 15919, but have a different meaning: (a) ISO 15919 uses "r̥" (ring below) for Sanskrit ऋ, and uses "ṛ" (dot below) for Hindi ड़. (b) ISO 15919 uses "ḷ" (ring below) for Sanskrit ऌ, and uses "ḷ" (dot below) for Vedic ळ. Sanskritists often use ṛ (dot below) for ऋ , ḷ (dot below) for ऌ and ḻ (macron below) for ळ, whereas ISO 15919 uses ḻ with macron below 1) for Tamil, 2) for ऴ (ळ with dot below). (c) ISO 15919 uses "ṃ" for "tippi" in Punjabi and "ṁ" for "anusvāra" in Sanskrit, Hindi, while many Sanskritists use "ṃ" and "ṁ" interchangeably for "anusvāra" in Sanskrit.

2. Non-ISO 15919 Diacritics

URW Palladio HOT also includes approx. 100 other diacritics, not defined or deprecated by ISO 15919, but used by many indologists, e.g. all the diacritics with "dot below" (Ṝ ṝ Ṛ ṛ etc.) instead of "ring below" (R̥̄ r̥̄ R̥ r̥ etc.), diacritics for the transliteration of anunāsika (Ḻ ḻ Ḷ ḷ), all diacritics for the transliteration of the Tibetan script, all diacritics used in TITUS files, etc. Furthermore, URW Palladio HOT includes all the diacritics contained in the superseded Unicode font URW Palladio UNI. For this reason, a few non-Indic diacritics, for instance Vietnamese Ơ ơ Ư ư, and Esperanto diacritics, are also contained in Palladio HOT to allow for completely replacing Palladio UNI by Palladio HOT being a superset of Palladio UNI.

Of the 572 glyphs of URW Palladio HOT, a total of approx. 300 diacritics are used for the transliteration of Indic scripts: approximately 200 ISO 15919 + 100 non-ISO 15919 diacritics. In the Glyph Lists (see below pages 17–21), all the ISO 15919 diacritics are marked with [1], and all the remaining non-ISO 15919 diacritics used by indologists are marked with [2].

3. How to use URW Palladio HOT

URW Palladio HOT is a hybrid OpenType font, which can be used either as an ordinary Unicode font (i.e. "Unicode mode") or as a featured OpenType font (i.e. "OpenType mode").

a) Non-Unicode-savvy Word Processor

If this is your case, you cannot use the font. You would have access to Code Page 1252 only.

b) Unicode-savvy Word Processor ("Unicode mode")

If this is your case, you can access all 572 glyphs, comprising the Unicode characters as well as the non-standard Private Use Area (PUA) characters. For the PUA characters, however, it is necessary to enter the PUA codes. For instance, in order to enter Ṛ, which has not been defined by the Unicode Consortium, you must enter the PUA code U+F120, either by a keyboard shortcut or by clicking on the character map. This is the backwards-compatible usage of this Hybrid OpenType font. When you upgrade to an OT-savvy word processor, you will be able to use the native OpenType features of URW Palladio HOT.

c) OpenType-savvy Word Processor ("OpenType mode")

If this is your case, you will never have to enter any codes for precomposed diacritical characters. You can enter all ISO 15919 and all non-ISO diacritics via the canonical stacking order. Example: In order to generate Ṛ, you would enter R•, which would be automatically displayed as Ṛ by the URW Palladio HOT internal OpenType glyph substitution routines. Text files created in this Opentype manner will never contain any Private Use Area codes.

As URW Palladio HOT is an innovative approach to solve the problems of diacritical fonts, you should be familiar with the basics of previous and present fonts and you should know how OpenType fonts are handled by the latest word processors available in 2004.

3.1. Types of Fonts

1. <u>PostScript Type 1</u> fonts address glyphs (outlines of characters) by <u>names</u>. For instance, glyph "a" is addressed by the name "a", "ˆ" by "circumflex", and "â" by "acircumflex".

2. <u>TrueType fonts</u> address glyphs by 4-digit hexadecimal Unicode <u>numbers</u>. For instance, glyph "a" is addressed by the hex number U+0061, "ˆ" by U+02C6, and "â" by U+00E2.

Even though PS and TT fonts can contain thousands of glyphs, in older word processors only 256 characters (32 control characters plus 224 visible characters) could be addressed, because the characters of text files were stored as "<u>bytes</u>" (i.e. 2-digit hexadecimal numbers).

Example: "ˆ" was <u>not</u> stored as hex 02C6, as you might think, but as hex 88 (decimal 136). Unicode double bytes were downsized to single bytes via Code Page Mapping (1252 etc.).

Since Word 1997, characters of text files may be stored as "<u>words</u>" (i.e. 4-digit hex numbers). Now hundreds or even thousands of glyphs contained in TrueType fonts are accessible. This is not true for PostScript Type 1 fonts, and therefore, a new font format was devised by Adobe and Microsoft: the OpenType font, which can be used both on PC and on Macintosh (hence the name "open": open to all operating systems) and which comes in 2 sub-formats:

1. <u>Extension ".ttf"</u>: OpenType font, internally a TT font (Basis splines, TT instructions, etc.)

2. <u>Extension ".otf"</u>: OpenType font, internally a PS font (Bézier curves, PS hints, etc.)

The <u>more conservative</u> format is the <u>TTF format</u> (file name extension ".ttf" as before), which may (or may not) contain additional internal tables for the handling of OpenType features. If a TT font contains no OT features, it is a plain TT font as before. But if a TT font contains OT features – implemented as GSUB (GS Glyph Substitutions) or GPOS (Glyph Positions) – it is a "featured" TT font. Those word processors that are incapable of handling OT features, however, will ignore these features and thus these fonts will be treated as if they did not contain any OT features ("Unicode mode"). On the other hand, word processors that are capable of handling OT features will use TT fonts as real OT fonts ("OpenType mode").

The <u>more innovative</u> format is the <u>OTF format</u> (file name extension ".otf"), which is also called compact font format (CFF). The greatest advantage of this format is its compactness: a TrueType font with a size of 300 KB shrinks to 100 KB, if converted to the OTF format. However, most word processors, printer drivers and pdf drivers cannot handle OTF fonts properly, especially if they contain large numbers of glyphs outside the range of the former PostScript Type 1 fonts with Adobe Standard Encoding. Therefore URW Palladio HOT is presently offered in the conservative TTF format only.

3.2. Unicode

The Unicode (Universal Code) Consortium allocated "code points" (hexadecimal numbers) to more than 50,000 characters. Nevertheless, many diacritics used by indologists will never be allocated code points by the Unicode Consortium.

The diacritics of the font URW Palladio HOT belong to four classes of Unicode characters:

1. <u>Base characters</u> are the simple letters a–z and A–Z (i.e. letters without diacritical marks) and a few other letters such as "ð", "ə", "ø", also regarded by Unicode as base characters.

2. <u>Combining diacritical marks</u> (◌̀ ◌́ ◌̂ ◌̃ ◌̄ ◌̣ ◌̤ ◌̥ ◌̦ ◌̧ etc.) are "accents", in that they are not used by themselves but needed to compose (typeset) composite characters (composites). A composite consists of a base character and one or more combining diacritical marks.

3. <u>Unicode-precomposed characters</u> are accented characters for which code points have been defined by Unicode, e.g. ẖ = U+1E96, Ḕ = U+1E14, Ṝ = U+1E5C, etc.

4. <u>Non-Unicode-precomposed characters</u> are accented characters for which code points have <u>not</u> been defined by Unicode, e.g. H̲, Ẽ, Ṟ, etc. Many diacritical characters defined by ISO 15919 and needed by indologists belong in this class. For them only "private" codes can be defined within the Private Use Area (U+E000 through U+F8FF).

3.3. Canonical Stacking Order

The Unicode Consortium expects that intelligent word processors be capable of composing <u>all</u> accented characters by <u>base</u> characters plus composing diacritical marks input via canonical stacking order. Therefore the Unicode Consortium does <u>not</u> intend to enlarge the number of <u>precomposed</u> characters, with the result that many accented characters used by indologists will never have official code points defined by the Unicode Consortium.

The prescribed sequence to input combinations of base character and combining diacritical marks is the canonical stacking order. It is illustrated by the following examples.

Stacking Order	R̥	R̥̄	R̥̄́	Ū	Ū́
3. Combining accent			◌́		◌́
2. Combining accent		◌̄	◌̄	◌̄	◌̄
0. Base Character	R	R	R	U	U
1. Combining accent	◌̥	◌̥	◌̥		
Left to right (0123):	R◌̥	R◌̥◌̄	R◌̥◌̄◌́	U◌̄	U◌̄◌́

<u>Note:</u> When inputting text, the base character always precedes combining diacritical marks. Since accented letters used by indologists have, at the most, one mark below and two marks above, more complicated stackings, e.g. with off-center accents, need not be explained here.

Canonical stacking has three major advantages:

1. Ease of Typing

All you need to know is a selection of these accents: ̀ ́ ̂ ̃ ̄ ̆ ̇ ̈ ̊ ̋ ̌ ̇ ̣ ̧ ̦ ̩ ̨ ̮,
plus the base character a–z and A–Z. No more consulting lists of hundreds of strange
hexadecimal codes for precombined diacritical characters! For instance, in order to key in
"ā", you need <u>not</u> know the code U+0101 for "ā", but you simply type "a" followed by "̄"
automatically resulting in "ā", provided you use an intelligent word processing program.
Hundreds of precomposed diacritical characters would require the definition of hundreds
of keyboard macros, while in order to follow the canonical stacking order you will have to
define only a few accent shortcuts, by which you can type hundreds of accented characters.
This is the greatest advantage of URW Palladio HOT: hundreds of precomposed diacritical
characters melt to a few accents combinable with virtually any base character (a-z and A-Z).

2. Ease of Sorting

If words with diacritical characters are entered in the canonical stacking order, they are
automatically sorted in the alphabetical a-to-z ascending order by Unicode hex numbers.
Example: "ā nanda" (0061, 0304 ...) is inserted after "ananda" (0061, 006E ...), whereas
"ānanda" (0101, 006E ...) is inserted after "z" (007A), since U+0101 is a larger number than
U+007A. It is apparent, however, that a series of words with diacritics entered in canonical
stacking order, and therefore sorted by the <u>Latin</u> alphabet, is not arranged in accord with
the <u>Sanskrit</u> alphabet. Nevertheless, this "Latinized sorting" is much more useful than the
"randomized sorting" that results, whenever precomposed diacritical letters occur in words.

3. Ease of Searching

Canonical stacking prescribed by the Unicode Consortium guarantees unambiguousness
(example: U+01D5 (U̇̈ ̄) and U+1E7A (U̇̄ ̈) are different code points). Canonical stacking
is also required for efficiently searching electronic files of indological texts. Searching those
files would be futile, however, if the canonical stacking order were not strictly adhered to.
Example: if, in an electronic file, ŕ in ŕsī was encoded partly by ṛ́ (ŕ = code U+0155), partly
by ṛ́ (ṛ = code U+1E5B), and partly by ṛ́ (ṛ́ = PUA code U+F119), you will never find all
occurrences of ṛ́sī, unless you waste your time trying out all non-canonical combinations.

Non-canonical stacking

Though Unicode allows for stacking of already stacked characters (i.e. <u>additional</u> stacking
of <u>precomposed</u> characters, so that both ṛ́ and ṛ́ are allowed by Unicode), this would
require the knowledge of all hex codes (U+0155 for ŕ and U+1E5B for ṛ). On the other hand,
to input ṛ́, you do not even need to know the code for "r", available on any keyboard.
URW Palladio HOT does not support the <u>non</u>-canonical stacking of already stacked letters,
because this would nullify the three major advantages of canonical stacking with regard to
(1) convenience of typing, (2) convenience of sorting and (3) convenience of searching.

3.4. Glyph Substitution

Of the 572 characters contained in URW Palladio HOT, 459 characters have been allocated code points by the Unicode Consortium. The remaining 113 characters (both ISO 15919 and non-ISO Indic diacritics) have never been, and apparently will never be, allocated official code points by the Unicode Consortium. It would have been possible to allocate glyph names only (without glyph codes) to these 113 glyphs, but in this case, URW Palladio HOT would have been a non-hybrid OpenType font for OpenType-savvy word processors only with the consequence that all users of non-OT-savvy word processors would not be able to access these 113 glyphs at all. Therefore these 113 glyphs were allocated PUA code points in range U+F100 through U+F1FF, a sub-range of the PUA range U+E000 through U+F8FF.

URW Palladio HOT contains about 350 precombined diacritical characters (Āā through Źź), and for all precombined letters (except a few special letters marked below with "No GS!"), OT Glyph Substitutions (GS) have been implemented in the OT font URW Palladio HOT.

Character codes and glyphs are two different things for OpenType-savvy word processors:

a) Non-OpenType-savvy word processor:

1. If you want to display the precomposed diacritical character "Ā", you must enter the code U+0100. This code will be stored and the respective glyph "Ā" will be displayed.

2. If you want to display the precomposed diacritical character "Ã", which has not been defined by Unicode, you must enter the font-internal PUA code U+F104. This PUA code will be stored in the text file and the respective PUA glyph "Ã" will be displayed.

b) OpenType-savvy word processor:

1. If you want to display the base character "A", you must hit the key "A". Its code U+0041, which you need not know, will be stored and the respective glyph "A" will be displayed.

2. If you want to display the precomposed diacritical character "Ā", you should press the key "A" (U+0041) followed by a macro for "◌̄" (U+0304). Codes U+0041 and U+0304 will be stored, but glyph "Ā", whose code U+0100 you need not know, will be displayed.

3. If you keep trying and enter a macro for "◌̃" (U+0303), this code will be stored, but the entire sequence "A◌̄◌̃" is displayed as glyph "Ã", whose font-internal PUA code U+F104 is not stored in the file. Now, if you continue and hit on the backspace key, "◌̃" (U+0303) will be deleted from storage and the remainder "A◌̄" will be displayed again as "Ā".

With OT-savvy word processors and with this "hot" OT-font URW Palladio HOT, the only code points you need to know are those of the accents: ◌̀ ◌́ ◌̂ ◌̃ ◌̄ ◌̇ ◌̈ ◌̊ ◌̋ ◌̆ ◌̣ ◌̤ ◌̥ ◌̦ ◌̱ As the accents ◌̇ ◌̋ ◌̦ ◌̥ ◌̱ etc. are rarely used by indologists, there remain ten-odd accents. Only for these accents, macros (keyboard shortcuts like Alt-m, Ctrl-z, or on the Macintosh Option-a etc.) have to be defined. No more is required. It's as simple as that! And the word processor files created in this way will never contain any Private Use Area character codes.

3.5. Word Processors

Before explaining how word processors handle OpenType fonts, the contents of the test file "pahot" is shown below. This test file is available for URW Palladio HOT in three versions: 1. pahot.doc (Word file), 2. pahot.txt (UTF-16 file), 3. pahot.htm (UTF-8 file).

<u>Note:</u> The test file is sorted by Unicode code points to demonstrate how this sorting works.

◌̀ 0300	A◌̃ = Ã	e◌́ = é	G◌̌ = Ğ
◌́ 0301	a◌̃ = ã	E◌̀ = È	g◌̌ = ğ
◌̂ 0302	A◌̊ = Å	e◌̀ = è	G◌̌ = Ğ
◌̃ 0303	a◌̊ = å	E◌̇ = Ė	g◌̌ = ğ
◌̄ 0304	A◌̨ = Ą	e◌̇ = ė	G◌̧ = Ģ
◌̆ 0306	a◌̨ = ą	E◌̂ = Ê	g◌̧ = ģ
◌̇ 0307	A◌̄◌́ = Ā́ *	e◌̂ = ê	H◌̂ = Ĥ
◌̈ 0308	a◌̄◌́ = ā́ *	E◌̈ = Ë	h◌̂ = ĥ
◌̊ 030A	A◌̄◌̀ = Ā̀ *	e◌̈ = ë	H◌̣ = Ḥ
◌̋ 030B	a◌̄◌̀ = ā̀ *	E◌̌ = Ě	h◌̣ = ḥ
◌̌ 030C	A◌̄◌̃ = Ā̃ *	e◌̌ = ě	H◌̣ = Ḥ
◌̐ 0310	a◌̄◌̃ = ā̃ *	E◌̌ = Ě	h◌̣ = ḥ
◌̣ 0323	A◌̄◌̆ = Ā̆ *	e◌̌ = ě	h◌̱ = ḥ
◌̤ 0324	a◌̄◌̆ = ā̆ *	E◌̄ = Ē	H◌̱ = Ḥ *
◌̥ 0325	A◌̆=Ă	e◌̄ = ē	H◌̤ = Ḥ *
◌̧ 0327	a◌̆=ă	E◌̃ = Ẽ	h◌̤ = ḥ *
◌̨ 0328	Æ◌̄ = Ǣ	e◌̃ = ẽ	I◌́ = Í
◌̮ 032E	æ◌̄ = ǣ	E◌̨ = Ę	i◌́ = í
◌̱ 0331	C◌́ = Ć	e◌̨ = ę	I◌̀ = Ì
	c◌́ = ć	E◌̄◌́ = Ḗ	i◌̀ = ì
	C◌̂ = Ĉ	e◌̄◌́ = ḗ	I◌̂ = Î
A◌́ = Á	c◌̂ = ĉ	E◌̄◌̀ = Ḕ	i◌̂ = î
a◌́ = á	C◌̌ = Č	e◌̄◌̀ = ḕ	I◌̈ = Ï
A◌̀ = À	c◌̌ = č	E◌̄◌̃ = Ẽ *	i◌̈ = ï
a◌̀ = à	C◌̧ = Ç	e◌̄◌̃ = ẽ *	I◌̆ = Ĭ
A◌̂ = Â	c◌̧ = ç	E◌̄◌̆ = Ĕ *	i◌̆ = ĭ
a◌̂ = â	D◌̌ = Ď	e◌̄◌̆ = ĕ *	I◌̄ = Ī
A◌̈ = Ä	d◌̌ = ď	fi = fi **	i◌̄ = ī
a◌̈ = ä	D◌̣ = Ḍ	fl = fl **	I◌̃ = Ĩ
A◌̆ = Ă	d◌̣ = ḍ	G◌̇ = Ġ	i◌̃ = ĩ
a◌̆ = ă	D◌̱ = Ḏ	g◌̇ = ġ	I◌̨ = Į
A◌̄ = Ā	d◌̱ = ḏ	G◌̂ = Ĝ	i◌̨ = į
a◌̄ = ā	E◌́ = É	g◌̂ = ĝ	I◌̄◌́ = Ī́ *

i꞉꞉ = í́ *	l꞉꞉ = l̥̄ *	O꞉ = Ŏ	r꞉ = r̥ *
I꞉꞉ = Ì̀ *	L꞉ = Ĺ *	o꞉ = ŏ	R꞉꞉ = R̥̄ *
i꞉꞉ = ì̀ *	l꞉ = ĺ *	O꞉ = Ō	r꞉꞉ = r̥̄ *
I꞉꞉ = Ĩ̃ *	L꞉ = L̀ *	o꞉ = ō	R꞉꞉ = Ŕ̥ *
i꞉꞉ = ĩ̃ *	l꞉ = l̀ *	O꞉ = Õ	r꞉꞉ = ŕ̥ *
I꞉꞉ = Ĭ̆ *	L꞉ = L̃ *	o꞉ = õ	R꞉꞉ = R̥̀ *
i꞉꞉ = ĭ̆ *	l꞉ = l̃ *	O꞉ = Ő	r꞉꞉ = r̥̀ *
I꞉ = Ĭ	L꞉ = Ľ *	o꞉ = ő	R꞉꞉꞉ = R̥̄́ *
i꞉ = ĭ	l꞉ = ľ *	O꞉꞉ = Ő	r꞉꞉꞉ = r̥̄́ *
Jˆ = Ĵ	M꞉ = Ṁ	o꞉꞉ = ő	R꞉꞉꞉ = R̥̄̀ *
jˆ = ĵ	m꞉ = ṁ	O꞉꞉ = Ò	r꞉꞉꞉ = r̥̄̀ *
jˇ = ǰ	M꞉ = M̥	o꞉꞉ = ò	R꞉ = Ř *
J꞉ = J́ *	m꞉ = m̥	O꞉꞉ = Õ *	r꞉ = ř *
j꞉ = j́ *	M꞉ = M̥̌ *	o꞉꞉ = õ *	R꞉ˆ = R̥̂ *
J꞉ = J̌ *	m꞉ = m̥̌ *	O꞉꞉ = Ŏ *	r꞉ˆ = r̥̂ *
K꞉ = Ķ	M꞉ = M̆ *	o꞉꞉ = ŏ *	S꞉ = Ś
k꞉ = ķ	m꞉ = m̆ *	O꞉ = Ŏ	s꞉ = ś
K꞉ = K̲	N꞉ = Ń	o꞉ = ŏ	Sˆ = Ŝ
k꞉ = k̲	n꞉ = ń	R꞉ = Ŕ	sˆ = ŝ
K↑H꞉ = K̲H *	N꞉ = Ṅ	r꞉ = ŕ	S꞉ = Š
K↑h꞉ = K̲h *	n꞉ = ṅ	R꞉ = Ṙ	s꞉ = š
k↑h꞉ = k̲h *	N꞉ = Ň	r꞉ = ṙ	S꞉ = Ş
L꞉ = Ĺ	n꞉ = ň	R꞉ = Ř	s꞉ = ş
l꞉ = ĺ	N꞉ = Ñ	r꞉ = ř	S꞉ = Ṣ
L꞉ = Ľ	n꞉ = ñ	R꞉ = Ṛ	s꞉ = ṣ
l꞉ = ľ	N꞉ = Ṇ	r꞉ = ṛ	S꞉ = S̲ *
L꞉ = Ļ	n꞉ = ṇ	R꞉ = Ṛ	s꞉ = s̲ *
l꞉ = ļ	N꞉ = N̥	r꞉ = ṛ	S꞉ = Ṣ̲ *
L꞉ = L̦	n꞉ = n̥	R꞉ = R̲	s꞉ = ṣ̲ *
l꞉ = l̦	N꞉ = N̲	r꞉ = r̲	T꞉ = Ť
L꞉ = L̲	n꞉ = n̲	R꞉꞉ = R̄	t꞉ = ť
l꞉ = l̲	N꞉ = Ň *	r꞉꞉ = r̄	T꞉ = Ţ
L꞉꞉ = L̄	n꞉ = ň *	R꞉꞉ = Ŕ̥ *	t꞉ = ţ
l꞉꞉ = l̄	O꞉ = Ó	r꞉꞉ = ŕ̥ *	T꞉ = Ṭ
L꞉꞉ = Ĺ *	o꞉ = ó	R꞉꞉ = R̥̀ *	t꞉ = ṭ
l꞉꞉ = ĺ *	O꞉ = Ò	r꞉꞉ = r̥̀ *	T꞉ = T̲
L꞉꞉ = L̀ *	o꞉ = ò	R꞉꞉꞉ = R̥̄́ *	t꞉ = t̲
l꞉꞉ = l̀ *	Oˆ = Ô	r꞉꞉꞉ = r̥̄́ *	T꞉ = Ṱ *
L꞉ = L̥ *	oˆ = ô	R꞉꞉꞉ = R̥̄̀ *	t꞉ = ṱ *
l꞉ = l̥ *	O꞉ = Ö	r꞉꞉꞉ = r̥̄̀ *	U꞉ = Ú
L꞉꞉ = L̥̄ *	o꞉ = ö	R꞉ = R̥̄ *	u꞉ = ú

U◌̀ = Ù	U◌́ = Ű	U◌̈◌̆=Ǚ	Z◌́ = Ź
u◌̀ = ù	u◌́ = ű	u◌̈◌̆=ǚ	z◌́ = ź
U◌̂ = Û	U◌̄◌́ = Ú *	U◌̈◌̄=Ǖ	Z◌̇ = Ż
u◌̂ = û	u◌̄◌́ = ú *	u◌̈◌̄=ǖ	z◌̇ = ż
U◌̈ = Ü	U◌̄◌̀ = Ù̄ *	Y◌́ = Ý	Z◌̌ = Ž
u◌̈ = ü	u◌̄◌̀ = ù̄ *	y◌́ = ý	z◌̌ = ž
U◌̆ = Ŭ	U◌̄◌̃ = Ũ̄ *	Y◌̀ = Ỳ	Z◌̣ = Ẓ
u◌̆ = ŭ	u◌̄◌̃ = ũ̄ *	y◌̀ = ỳ	z◌̣ = ẓ
U◌̄ = Ū	U◌̄◌̆ = Ŭ̄ *	Y◌̇ = Ẏ	Z◌̱ = Z̲
u◌̄ = ū	u◌̄◌̆ = ŭ̄ *	y◌̇ = ẏ	z◌̱ = z̲
U◌̃ = Ũ	U◌̌=Ǔ	Y◌̂ = Ŷ	Z◌̤ = Ẕ *
u◌̃ = ũ	u◌̌=ǔ	y◌̂ = ŷ	z◌̤ = ẕ *
U◌̊ = Ů	U◌̈◌́=Ǘ	y◌̈ = ÿ	
u◌̊ = ů	u◌̈◌́=ǘ	Y◌̈ = Ÿ	* PUA GSUB
U◌̨ = Ų	U◌̈◌̀=Ǜ	Y◌̃ = Ỹ	** Adobe GSUB
u◌̨ = ų	u◌̈◌̀=ǜ	y◌̃ = ỹ	

1. On the <u>left</u> of "=" all the diacritical characters, for which glyph substitutions have been implemented in the font URW Palladio HOT are stored in the <u>canonical stacking order.</u> E.g. "A◌̄" is stored in the test file "pahot" internally as U+0041 U+0304.

2. On the <u>right</u> of "=" all these diacritical characters are stored as <u>precomposed characters,</u> which are also available in URW Palladio HOT. E.g. "Ā" is stored as U+0100 in "pahot".

When you open/import "pahot" (.doc .txt .htm) into your word processor, this may happen:

1. If on the <u>right</u> of "=" you see precomposed characters only, then your word processor is fully Unicode-savvy, i.e. it can handle both PUA and non-PUA precomposed characters.

2. If on the <u>left</u> of "=" you see precomposed characters only, e.g. if you see "Ā = Ā" instead of "A◌̄ = Ā" etc., then your word processor is fully OpenType-savvy.

3. If on the <u>left</u> of the equations <u>not marked</u> with "*" you see the same characters as on the right of "=", e.g. "Ā = Ā" etc., <u>and</u> if on the <u>left</u> of the equations <u>marked</u> with "*" you see the file-internal canonical stacking order, e.g. "R◌̥ = R̥ *", your word processor is <u>not</u> OpenType-savvy but can handle the canonical stacking for exactly those precomposed characters that have been allocated code points by the Unicode Consortium.

<u>Experiment:</u> Microsoft claims that Word XP <u>is</u> OpenType-savvy. So, start Word XP, enter "fl" and mark it up with the font "Palatino Linotype" supplied with Windows XP. You see: "fl" is <u>not</u> substituted by the ligature "fl". Result: Word XP <u>is not</u> OpenType-savvy.

The various approaches by word processors to OpenType fonts are as follows:

1. <u>Adobe InDesign</u> evaluates the glyph substitutions (GSUBs) in OpenType fonts for Latin scripts directly (without recourse to Uniscribe in the case of the Windows versions of InDesign), and therefore <u>all</u> GSUBs of URW Palladio HOT are correctly handled. When you load "pahot" into InDesign (tests were performed with InDesign 1.0 and 2.0 for Windows and with InDesign 2.0 for Macintosh), you will see on the left of "=" exactly the same as on the right of "=". Even the so-called "zero-width-joiner" ligature "K⇡h•" is correctly displayed as "<u>Kh</u>", and, of course, "fl" is correctly converted to the ligature "fl".

2. <u>Word 1997 and 2000</u> are Unicode-savvy and correctly display all precomposed diacritics to the right of "=", even those diacritics, for which PUA codes had to be defined, but Word 1997 and 2000 are neither OpenType-savvy nor able to handle canonical stacking, and you will see on the left of "=" the file-internal canonical stacking, i.e. "A•" as "A•".

3. <u>Word XP (Word 2002)</u> is not OpenType-savvy as far as OT fonts for Latin scripts are concerned. It completely ignores all OT-features contained in OT fonts for Latin scripts and treats these fonts, as if they were plain Unicode fonts. But Word XP has a built-in routine that converts the canonical stacking of exactly those precomposed letters, for which code points were allocated by Unicode (so-called "legacy diacritical characters"), so that you see "A•" as "Ā", whereas you see "R•" as "R•", because "Ṛ" does not exist for Unicode and hence it does not exist for Word XP. However, this canonical stacking by Word XP has a few bugs: "G•" is not displayed as "Ǧ" (while "g•" is displayed as "ǧ"). The handling of Ò ò Ó ó is buggy too: "O••" is displayed as "Ō•", "o••" is displayed as "Ò" (uppercase!), "O••" is displayed as "Ō•", and "o••" is displayed as "Ó" (uppercase!).

4. <u>Internet Explorer 6</u>: This brower for Windows uses exactly the same canonical stacking routine as Word XP, as is proved by exactly the same bugs, which have just been listed.

5. <u>TextEdit</u> (for Macintosh OS X) is not OpenType-savvy, but has a built-in routine that converts canonical stacking to exactly the same extent as Word XP. But the canonical stacking routine of TextEdit is flawless, so that the above-mentioned bugs of Word XP do not occur. Furthermore, TextEdit tries to stack those precomposed letters, for which code points were not allocated by the Unicode Consortium, by treating the combining diacritical marks (range U+300) as "floating accents", irrespective of whether these have been designed as zero-width deadkey accents or spacing accents. (This accent handling by TextEdit is similar to the accent handling by WordPerfect. But since it is not known, to what extent the so-called glyph bounding box is evaluated by TextEdit, no effort has been made to design a special "customized" Macintosh version of URW Palladio HOT that would enable TextEdit to <u>exactly</u> fit the floating accents of non-Unicode diacritics.)

The poor performance of word processsors with respect to OpenType fonts is a shame for the entire software industry, which is unable to supply OT-savvy programs, although the specifications for OpenType fonts have been published in 1997 – long five years ago. See: "OpenType Horror Picture Show of Word Processors" below on the pages 44–48.

3.6. Recommendations

Presently, only InDesign is capable of fully handling OT fonts for Latin scripts. The reason that Word XP does not support OT features (GSUB, GPOS) for Latin scripts is Uniscribe:

This dynamic link library <u>Uniscribe</u> (USP10.DLL = "Unicode Script Processor") supplied with Windows XP and Internet Explorer has been designed by Microsoft for "complex" scripts, e.g. for the Arabic and Hebrew scripts etc. It seems that Microsoft is of the opinion, that the Latin script is "no complex" script. Hence Microsoft offers no OpenType support for OpenType fonts with Latin scripts. Hence Word XP, which internally calls the Uniscribe routines, treats an OT font for Latin scripts (e.g. URW Palladio HOT), as if it were a usual Unicode font without any OpenType features. Why offers Microsoft "Linotype Palatino" as an OpenType font, although none of its OpenType features work with Microsoft Windows?

Word XP, and also Internet Explorer 6, can only handle the canonical stacking of exactly those precomposed characters of the Latin script (with code points in range up to U+1EFF), for which code points were allocated by the Unicode Consortium. Word XP and IE 6 handle the canonical stacking directly by their own tables (without recourse to GSUBs in OT fonts).

The precomposed characters, listed below on pages 24–27 for URW Palladio HOT, comprise a total of 459 Unicode-allocated (i.e. non-PUA) characters. Now, the beauty of Unicode is that the code points for precomposed, non-PUA, characters are exactly the same for all Unicode-compliant fonts in all computer platforms. My tables on pages 24-27, therefore, are equally applicable to URW Palladio HOT as to any other Unicode font.

1. Those indologists who are willing to restrict themselves to the 459 characters listed on pages 24–27 can use URW Palladio HOT already now with Word XP and with Internet Explorer 6 and can offer text files, input via canonical stacking, on the internet for online browsing – provided the users have previously installed the font URW Palladio HOT. Such text files will be in accord with Unicode ("Stacking mode" for legacy diacritics). This implies, however, that you <u>may not</u> use several ISO 15919 diacritics, e.g. ā́, r̥, l̥, etc.

2. Those indologists who are unwilling to restrict themselves to those 459 characters listed on pages 24–27, and who want to use some or all of the other PUA diacritical characters listed on page 27, will have to use InDesign as their word processor for the time being ("OpenType mode"). As regards the internet, they will have to develop their own browser, as it is unlikely that any browser will ever support canonical stacking of PUA characters that have never been allocated code points by the Unicode Consortium.

3. Those indologists who use older word processors such as Word 1997, Word 2000 etc. have no other option but to always enter <u>precomposed</u> diacritical characters directly, which necessitates the definition of hundreds of keyboard shortcuts ("Unicode mode"). They would be better off by using any of my various classic TrueType and PostScript Type 1 fonts, e.g. the non-Unicode fonts URW Palladio CSX+, URW Palladio SKT, etc.

4. Glyph Lists of URW Palladio HOT

4.1. Sorted Glyph List of Alphabetical Characters

[1] = ISO 15919 Indic Diacritics, [2] = Non-ISO 15919 Indic Diacritics

A	0041	65 [1]		AĪ	F160	61792 [2] (no GS!)		Ḍ	1E0C	7692 [1]
a	0061	97 [1]		aĩ	F161	61793 [2] (no GS!)		ḍ	1E0D	7693 [1]
Ā	0100	256 [1]		AŪ	F162	61794 [2] (no GS!)		Ḏ	1E0E	7694 [2]
ā	0101	257 [1]		aũ	F163	61795 [2] (no GS!)		ḏ	1E0F	7695 [2]
Ă	0102	258 [2]		Ǎ	01CD	461 [2]		Ď	010E	270 (caron!)
ă	0103	259 [2]		ǎ	01CE	462 [2]		ď	010F	271 (caron!)
À	00C0	192 [1]		Å	00C5	197				
à	00E0	224 [1]		å	00E5	229		Đ	00D0	208 (no GS!)
Á	00C1	193 [1]		Ą	0104	260		ð	00F0	240 (no GS!)
á	00E1	225 [1]		ą	0105	261		Đ	0110	272 (no GS!)
Ã	00C3	195 [1]						đ	0111	273 (no GS!)
ã	00E3	227 [1]		B	0042	66 [1]				
Â	00C2	194 [2]		b	0062	98 [1]		E	0045	69 [1]
â	00E2	226 [2]						e	0065	101 [1]
Ä	00C4	196 [2]		C	0043	67 [1]		Ē	0112	274 [1]
ä	00E4	228 [2]		c	0063	99 [1]		ē	0113	275 [1]
Ă	F106	61702 [2]		Ĉ	0108	264 [1]		Ĕ	0114	276 [2]
ă	F107	61703 [2]		ĉ	0109	265 [1]		ĕ	0115	277 [2]
À	F102	61698 [1]		Ç	00C7	199 [2]		È	00C8	200 [1]
à	F103	61699 [1]		ç	00E7	231 [2]		è	00E8	232 [1]
Á	F100	61696 [1]		Ć	0106	262		É	00C9	201 [1]
á	F101	61697 [1]		ć	0107	263		é	00E9	233 [1]
Ã	F104	61700 [1]		Č	010C	268		Ẽ	1EBC	7868 [1]
ã	F105	61701 [1]		č	010D	269		ẽ	1EBD	7869 [1]
Æ	00C6	198 [1] (no GS!)						Ê	00CA	202 [1]
æ	00E6	230 [1] (no GS!)						ê	00EA	234 [1]
Ǽ	01E2	482 [1] (via GS!)		D	0044	68 [1]		Ë	00CB	203 [2]
ǽ	01E3	483 [1] (via GS!)		d	0064	100 [1]		ë	00EB	235 [2]

Ě	F13A	61754 [2]	H	0048	72 [1]	Ì̄	F10A	61706 [1]	
ě	F13B	61755 [2]	h	0068	104 [1]	ì̄	F10B	61707 [1]	
Ḕ	1E14	7700 [1]	ẖ	1E96	7830 [1]	Í̄	F108	61704 [1]	
ḕ	1E15	7701 [1]	H̱	F14B	61771 [1]	í̄	F109	61705 [1]	
Ḗ	1E16	7702 [1]	Ḫ	1E2A	7722 [1]	Ĩ̄	F10C	61708 [1]	
ḗ	1E17	7703 [1]	ḫ	1E2B	7723 [1]	ĩ̄	F10D	61709 [1]	
Ē̃	F138	61752 [1]	Ḥ	1E24	7716 [1]	Ǐ	01CF	463 [2]	
ē̃	F139	61753 [1]	ḥ	1E25	7717 [1]	ǐ	01D0	464 [2]	
Ě	011A	282 [2]	Ḥ̱	F14C	61772 [1]	Į	012E	302	
ě	011B	283 [2]	ḥ̱	F14D	61773 [1]	į	012F	303	
Ǝ	018E	398 [2] (no GS!)	ᴴ	F15F	61791 [2] (no GS!)	İ	0130	304 (no GS!)	
ə	0259	601 [2] (no GS!)	ʰ	02B0	688 [2] (no GS!)	ı	0131	305 (no GS!)	
Ė	0116	278	Ĥ	0124	292				
ė	0117	279	ĥ	0125	293				
Ę	0118	280				J	004A	74 [1]	
ę	0119	281				j	006A	106 [1]	
			I	0049	73 [1]	J̌	F15A	61786 [2]	
F	0046	70 [1]	i	0069	105 [1]	ǰ	01F0	496 [2]	
f	0066	102 [1]	Ī	012A	298 [1]	J́	F158	61784 [2]	
			ī	012B	299 [1]	j́	F159	61785 [2]	
			Ĭ	012C	300 [2]	Ĵ	0134	308	
			ĭ	012D	301 [2]	ĵ	0135	309	
G	0047	71 [1]	Ì	00CC	204 [1]	ȷ	F6BE	63166 (no GS!)	
g	0067	103 [1]	ì	00EC	236 [1]				
Ġ	0120	288 [1]	Í	00CD	205 [1]				
ġ	0121	289 [1]	í	00ED	237 [1]	K	004B	75 [1]	
Ǧ	01E6	486 [2]	Ĩ	0128	296 [1]	k	006B	107 [1]	
ǧ	01E7	487 [2]	ĩ	0129	297 [1]	Ḵ	1E34	7732 [1]	
Ğ	011E	286	Î	00CE	206 [2]	ḵ	1E35	7733 [1]	
ğ	011F	287	î	00EE	238 [2]	Ḵh	F147	61767 [1] (ZWJ!)	
Ĝ	011C	284	Ï	00CF	207 [2]	ḵh	F148	61768 [1] (ZWJ!)	
ĝ	011D	285	ï	00EF	239 [2]	ḴH	F146	61766 [1] (ZWJ!)	
Ģ	0122	290 (cedilla!)	Ǐ	F10E	61710 [2]	Ķ	0136	310 (cedilla!)	
ģ	0123	291 (cedilla!)	ǐ	F10F	61711 [2]	ķ	0137	311 (cedilla!)	

Glyph	Code	Number	Note
L	004C	76	[1]
l	006C	108	[1]
Ḻ	1E3A	7738	[1]
ḻ	1E3B	7739	[1]
Ḷ	1E36	7734	[1] [2] (cave!)
ḷ	1E37	7735	[1] [2] (cave!)
Ḻ	F130	61744	[1]
ḻ	F131	61745	[1]
Ḹ	1E38	7736	[2]
ḹ	1E39	7737	[2]
Ḹ	F132	61746	[1]
ḹ	F133	61747	[1]
L̀	F12E	61742	[2]
l̀	F12F	61743	[2]
L̥	F136	61750	[1]
l̥	F137	61751	[1]
Ĺ	F12C	61740	[2]
ĺ	F12D	61741	[2]
Ĺ	F134	61748	[1]
ĺ	F135	61749	[1]
L̃	F142	61762	[2]
l̃	F143	61763	[2]
Ľ	F144	61764	[2]
ľ	F145	61765	[2]
Ĺ	0139	313	
ĺ	013A	314	
Ļ	013B	315	(cedilla!)
ļ	013C	316	(cedilla!)
Ľ	013D	317	(caron!)
ľ	013E	318	(caron!)
Ł	0141	321	(no GS!)
ł	0142	322	(no GS!)

Glyph	Code	Number	Note
M	004D	77	[1]
m	006D	109	[1]
Ṁ	1E40	7744	[1]
ṁ	1E41	7745	[1]
Ṃ	1E42	7746	[1]
ṃ	1E43	7747	[1]
M̌	F140	61760	[1]
m̌	F141	61761	[1]
M̆	F152	61778	[1]
m̆	F153	61779	[1]
N	004E	78	[1]
n	006E	110	[1]
N̲	1E48	7752	[1]
n̲	1E49	7753	[1]
Ň	F154	61780	[1]
ň	F155	61781	[1]
Ñ	00D1	209	[1]
ñ	00F1	241	[1]
Ṅ	1E44	7748	[1]
ṅ	1E45	7749	[1]
Ṇ	1E46	7750	[1]
ṇ	1E47	7751	[1]
Ń	0143	323	
ń	0144	324	
Ň	0147	327	
ň	0148	328	
Ņ	0145	325	(cedilla!)
ņ	0146	326	(cedilla!)
O	004F	79	[1]
o	006F	111	[1]
Ō	014C	332	[1]
ō	014D	333	[1]

Glyph	Code	Number	Note
Ŏ	014E	334	[2]
ŏ	014F	335	[2]
Ò	00D2	210	[1]
ò	00F2	242	[1]
Ó	00D3	211	[1]
ó	00F3	243	[1]
Õ	00D5	213	[1]
õ	00F5	245	[1]
Ô	00D4	212	[1]
ô	00F4	244	[1]
Ö	00D6	214	[2]
ö	00F6	246	[2]
Ŏ	F13E	61758	[2]
ŏ	F13F	61759	[2]
Ò	1E50	7760	[1]
ò	1E51	7761	[1]
Ó	1E52	7762	[1]
ó	1E53	7763	[1]
Õ	F13C	61756	[1]
õ	F13D	61757	[1]
Ǒ	01D1	465	[2]
ǒ	01D2	466	[2]
Ő	0150	336	
ő	0151	337	
Ơ	01A0	416	(no GS!)
ơ	01A1	417	(no GS!)
Ø	00D8	216	(no GS!)
ø	00F8	248	(no GS!)
Ɵ	019F	415	(no GS!)
ɵ	0275	629	(no GS!)
Œ	0152	338	(no GS!)
œ	0153	339	(no GS!)

P	0050	80 [1]	Ŕ̥	F11C	61724 [2]	T	0054	84 [1]
p	0070	112 [1]	ŕ̥	F11D	61725 [2]	t	0074	116 [1]
			Ř̥	F128	61736 [1]	Ṯ	1E6E	7790 [1]
Q	0051	81 [1]	ř̥	F129	61737 [1]	ṯ	1E6F	7791 [1]
q	0071	113 [1]	R̥̂	F15D	61789 [2]	Ṭ	1E6C	7788 [1]
			r̥̂	F15E	61790 [2]	ṭ	1E6D	7789 [1]
R	0052	82 [1]	Ŕ	0154	340	Ṭ̱	F150	61776 [1]
r	0072	114 [1]	ŕ	0155	341	ṭ̱	F151	61777 [1]
Ṟ	1E5E	7774 [1]	Ř	0158	344	Ţ	0162	354 (cedilla!)
ṟ	1E5F	7775 [1]	ř	0159	345	ţ	0163	355 (cedilla!)
Ř	F156	61782 [1]	Ŗ	0156	342 (cedilla!)	Ť	0164	356 (caron!)
ř	F157	61783 [1]	ŗ	0157	343 (cedilla!)	ť	0165	357 (caron!)
Ṙ	1E58	7768 [2]						
ṙ	1E59	7769 [2]	S	0053	83 [1]	Þ	00DE	222 (no GS!)
Ṛ	1E5A	7770 [1] [2] (cave!)	s	0073	115 [1]	þ	00FE	254 (no GS!)
ṛ	1E5B	7771 [1] [2] (cave!)	Ś	015A	346 [1]			
Ṛ̥	F120	61728 [1]	ś	015B	347 [1]	U	0055	85 [1]
ṛ̥	F121	61729 [1]	Ṣ	1E62	7778 [1]	u	0075	117 [1]
R̤	1E5C	7772 [2]	ṣ	1E63	7779 [1]	Ū	016A	362 [1]
r̤	1E5D	7773 [2]	Ṣ̱	F14E	61774 [1]	ū	016B	363 [1]
R̤̥	F122	61730 [1]	ṣ̱	F14F	61775 [1]	Ŭ	016C	364 [1]
r̤̥	F123	61731 [1]	S̱	F149	61769 [1]	ŭ	016D	365 [1]
Ȑ	F11A	61722 [2]	s̱	F14A	61770 [1]	Ù	00D9	217 [1]
ȑ	F11B	61723 [2]	Ŝ	015C	348	ù	00F9	249 [1]
Ȑ̥	F126	61734 [1]	ŝ	015D	349	Ú	00DA	218 [1]
ȑ̥	F127	61735 [1]	Š	0160	352	ú	00FA	250 [1]
Ŕ̠	F118	61720 [2]	š	0161	353	Ũ	0168	360 [1]
ŕ̠	F119	61721 [2]	Ş	015E	350 (true cedilla!)	ũ	0169	361 [1]
Ŕ̥	F124	61732 [1]	ş	015F	351 (true cedilla!)	Û	00DB	219 [2]
ŕ̥	F125	61733 [1]	Ș	0218	536 (no GS!)	û	00FB	251 [2]
R̥̀	F11E	61726 [2]	ș	0219	537 (no GS!)	Ü	00DC	220 [2]
r̥̀	F11F	61727 [2]				ü	00FC	252 [2]
R̥̀	F12A	61738 [1]	ß	00DF	223 (no GS!)	Ŭ	F116	61718 [2]
r̥̀	F12B	61739 [1]				ŭ	F117	61719 [2]

Ȕ	F112	61714 [1]		Ư	01AF	431 (no GS!)		Ŷ	0176	374
ȕ	F113	61715 [1]		ư	01B0	432 (no GS!)		ŷ	0177	375
Ű	F110	61712 [1]						Ÿ	0178	376
ű	F111	61713 [1]		V	0056	86 [1]		ÿ	00FF	255
Ȗ	F114	61716 [1]		v	0076	118 [1]				
ȗ	F115	61717 [1]						Z	005A	90 [1]
Ǔ	01D3	467 [2]		W	0057	87 [1]		z	007A	122 [1]
ǔ	01D4	468 [2]		w	0077	119 [1]		Ẕ	1E94	7828 [1]
Ǖ	01D5	469 [2]						ẕ	1E95	7829 [1]
ǖ	01D6	470 [2]		X	0058	88		Ż	017B	379 [1]
Ǘ	01D7	471 [2]		x	0078	120		ż	017C	380 [1]
ǘ	01D8	472 [2]						Ẓ	1E92	7826 [1]
Ǚ	01D9	473 [2]		Y	0059	89 [1]		ẓ	1E93	7827 [1]
ǚ	01DA	474 [2]		y	0079	121 [1]		Ẕ	F15B	61787 [2]
Ǜ	01DB	475 [2]		Ỹ	1EF8	7928 [2]		ẓ	F15C	61788 [2]
ǜ	01DC	476 [2]		ỹ	1EF9	7929 [2]		Ž	017D	381 [1]
Ů	016E	366		Ẏ	1E8E	7822 [1]		ž	017E	382 [1]
ů	016F	367		ẏ	1E8F	7823 [1]		Ź	0179	377
Ų	0172	370		Ỳ	1EF2	7922		ź	017A	378
ų	0173	371		ỳ	1EF3	7923				
Ű	0170	368		Ý	00DD	221				
ű	0171	369		ý	00FD	253				

4.2. Classified Glyph List of Figures, Symbols, Accents, etc.

Spaces:

	0020	32 (space)
	00A0	160 (nobreak s.)
‍	200D	8205 (ZWJ)

Figures:

0	0030	48
1	0031	49
2	0032	50
3	0033	51
4	0034	52
5	0035	53
6	0036	54
7	0037	55
8	0038	56
9	0039	57
¼	00BC	188
½	00BD	189
¾	00BE	190

Punctuation Marks:

,	002C	44 (comma)
;	003B	59 (semicolon)
.	002E	46 (fullstop)
:	003A	58 (colon)
'	0027	39 (apostrophe)
?	003F	63 (question)
¿	00BF	191 (Spanish q.)
!	0021	33 (exclamation)
¡	00A1	161 (Spanish e.)
*	002A	42 (asterisk)
†	2020	8224 (dagger)
‡	2021	8225 (double d.)
·	00B7	183 (middle dot)
•	2022	8226 (bullet)

…	2026	8230 (ellipsis)
§	00A7	167 (section)
¶	00B6	182 (paragraph)

~	007E	126 (ascii tilde)
^	005E	94 (ascii circum)

Brackets:

(	0028	40
)	0029	41
[	005B	91
]	005D	93
{	007B	123
}	007D	125

Quotation Marks:

'	2018	8216 (single left)
'	2019	8217 (single right)
,	201A	8218 (single base)

"	0022	34 (quote & inch)
"	201C	8220 (double left)
"	201D	8221 (d. right)
„	201E	8222 (d. base)

‹	2039	8249 (single left)
›	203A	8250 (single right)
«	00AB	171 (double left)
»	00BB	187 (double right)

Ligatures:

&	0026	38 (et ligature)
@	0040	64 (at ligature)
fi	FB01	64257 (via GS)
fl	FB02	64258 (via GS)

Slashes:

/	002F	47 (slash)
\	005C	92 (backslash)
|	007C	124 (bar)
¦	00A6	166 (broken bar)

Hyphens:

_	005F	95 (underscore)
-	002D	45 (hard hyphen)
-	00AD	173 (soft hyphen)
–	2013	8211 (en dash)
—	2014	8212 (em dash)

Currency:

$	0024	36 (Dollar)
¢	00A2	162 (Cent)
£	00A3	163 (Sterling)
ƒ	0192	402 (Florin)
€	20AC	8364 (Euro)
¥	00A5	165 (Yen)
¤	00A4	164 (currency)

Copyright:

©	00A9	169 (copyright)
®	00AE	174 (registered)
™	2122	8482 (trademark)

Superior Figures:

[1]	00B9	185
[2]	00B2	178
[3]	00B3	179

[a]	00AA	170 (Spanish Nª)
°	00B0	176 (Spanish N°)
º	00BA	186 (degree)

Column 1

Mathematical Symbols:

+	002B	43 (plus)
−	2212	8722 (minus)
×	00D7	215 (multiply)
÷	00F7	247 (divide)
⁄	2044	8260 (fraction)
¬	00AC	172 (logical not)
±	00B1	177 (plusminus)
#	0023	35 (numbersign)
<	003C	60
>	003E	62
=	003D	61
≤	2264	8804
≥	2265	8805
≠	2260	8800
%	0025	37
‰	2030	8240
Δ	2206	8710 (delta)
◊	25CA	9674 (lozenge)
Σ	2211	8721 (summation)
√	221A	8730 (radical)
∂	2202	8706 (partialdiff)
μ	00B5	181 (mu)

Apple Mac OS Symbols:

∞	221E	8734 (infinity)
∏	220F	8719 (product)
π	03C0	960 (pi)
∫	222B	8747 (integral)

Column 2

Ω	2126	8486 (Omega)
≈	2248	8776 (approx)
	F8FF	63743 (logo)

Metrics Symbols:

˘	F1F8	61944 (short)
–	F1F9	61945 (long)
᷑	F1FA	61946 (sh. or l.)
᷒	F1FB	61947
ᷓ	F1FC	61948
ᷔ	F1FD	61949
ᷕ	F1FE	61950
ᷖ	F1FF	61951

Spacing Accents:

`	0060	96 (grave)
´	00B4	180 (acute)
^	02C6	710 (circumflex)
˜	02DC	732 (tilde)
¯	00AF	175 (macron)
˘	02D8	728 (breve)
˙	02D9	729 (dotaccent)
¨	00A8	168 (dieresis)
°	02DA	730 (ring)
˝	02DD	733 (hungaruml.)
ˇ	02C7	711 (caron)
¸	00B8	184 (cedilla)
˛	02DB	731 (ogonek)
'	F6C3	63171 (comma-a.)

Column 3

Arabic Signs:

ʾ	02BE	702 (ringhalf-r)
ʿ	02BF	703 (ringhalf-left)

Cave! No apostrophes, but Arabic hamza (ʾ) and ain (ʿ)

Combining Accents:

̀	0300	768 (grave)
́	0301	769 (acute)
̂	0302	770 (circumflex)
̃	0303	771 (tilde)
̄	0304	772 (macron)
̆	0306	774 (breve)
̇	0307	775 (dotaccent)
̈	0308	776 (dieresis)
̊	030A	778 (ring)
̋	030B	779 (hungaruml.)
̌	030C	780 (caron)
̐	0310	784 (candrabin.)
̣	0323	803 (dot below)
̤	0324	804 (dieresis b.)
̥	0325	805 (ring below)
̧	0327	807 (cedilla)
̨	0328	808 (ogonek)
̮	032E	814 (breve below)
̱	0331	817 (macron b.)

Cave! The avagraha (as in "ko 'pi" = को ऽपि) is usually encoded as apostrophe proper (0027), but TITUS files use "modifier apostrophe" (02BC) and even Armenian apostrophe (055A), while ISO 15919 recommends the British-American single left quotation mark (2018).

	0020	32	C	0043	67	f	0066	102	ª	00AA	170
!	0021	33	D	0044	68	g	0067	103	«	00AB	171
"	0022	34	E	0045	69	h	0068	104	¬	00AC	172
#	0023	35	F	0046	70	i	0069	105	-	00AD	173
$	0024	36	G	0047	71	j	006A	106	®	00AE	174
%	0025	37	H	0048	72	k	006B	107	¯	00AF	175
&	0026	38	I	0049	73	l	006C	108	°	00B0	176
'	0027	39	J	004A	74	m	006D	109	±	00B1	177
(	0028	40	K	004B	75	n	006E	110	²	00B2	178
)	0029	41	L	004C	76	o	006F	111	³	00B3	179
*	002A	42	M	004D	77	p	0070	112	´	00B4	180
+	002B	43	N	004E	78	q	0071	113	µ	00B5	181
,	002C	44	O	004F	79	r	0072	114	¶	00B6	182
-	002D	45	P	0050	80	s	0073	115	·	00B7	183
.	002E	46	Q	0051	81	t	0074	116	¸	00B8	184
/	002F	47	R	0052	82	u	0075	117	¹	00B9	185
0	0030	48	S	0053	83	v	0076	118	º	00BA	186
1	0031	49	T	0054	84	w	0077	119	»	00BB	187
2	0032	50	U	0055	85	x	0078	120	¼	00BC	188
3	0033	51	V	0056	86	y	0079	121	½	00BD	189
4	0034	52	W	0057	87	z	007A	122	¾	00BE	190
5	0035	53	X	0058	88	{	007B	123	¿	00BF	191
6	0036	54	Y	0059	89			007C	124	À	00C0	192
7	0037	55	Z	005A	90	}	007D	125	Á	00C1	193
8	0038	56	[	005B	91	~	007E	126	Â	00C2	194
9	0039	57	\	005C	92		00A0	160	Ã	00C3	195
:	003A	58	]	005D	93	¡	00A1	161	Ä	00C4	196
;	003B	59	^	005E	94	¢	00A2	162	Å	00C5	197
<	003C	60	_	005F	95	£	00A3	163	Æ	00C6	198
=	003D	61	`	0060	96	¤	00A4	164	Ç	00C7	199
>	003E	62	a	0061	97	¥	00A5	165	È	00C8	200
?	003F	63	b	0062	98	¦	00A6	166	É	00C9	201
@	0040	64	c	0063	99	§	00A7	167	Ê	00CA	202
A	0041	65	d	0064	100	¨	00A8	168	Ë	00CB	203
B	0042	66	e	0065	101	©	00A9	169	Ì	00CC	204

Í	00CD	205	ò	00F2	242	ę	0119	281	Ņ	0145	325
Î	00CE	206	ó	00F3	243	Ě	011A	282	ņ	0146	326
Ï	00CF	207	ô	00F4	244	ě	011B	283	Ň	0147	327
Ð	00D0	208	õ	00F5	245	Ĝ	011C	284	ň	0148	328
Ñ	00D1	209	ö	00F6	246	ĝ	011D	285	Ō	014C	332
Ò	00D2	210	÷	00F7	247	Ğ	011E	286	ō	014D	333
Ó	00D3	211	ø	00F8	248	ğ	011F	287	Ŏ	014E	334
Ô	00D4	212	ù	00F9	249	Ġ	0120	288	ŏ	014F	335
Õ	00D5	213	ú	00FA	250	ġ	0121	289	Ő	0150	336
Ö	00D6	214	û	00FB	251	Ģ	0122	290	ő	0151	337
×	00D7	215	ü	00FC	252	ģ	0123	291	Œ	0152	338
Ø	00D8	216	ý	00FD	253	Ĥ	0124	292	œ	0153	339
Ù	00D9	217	þ	00FE	254	ĥ	0125	293	Ŕ	0154	340
Ú	00DA	218	ÿ	00FF	255	Ĩ	0128	296	ŕ	0155	341
Û	00DB	219	Ā	0100	256	ĩ	0129	297	Ŗ	0156	342
Ü	00DC	220	ā	0101	257	Ī	012A	298	ŗ	0157	343
Ý	00DD	221	Ă	0102	258	ī	012B	299	Ř	0158	344
Þ	00DE	222	ă	0103	259	Ĭ	012C	300	ř	0159	345
ß	00DF	223	Ą	0104	260	ĭ	012D	301	Ś	015A	346
à	00E0	224	ą	0105	261	Į	012E	302	ś	015B	347
á	00E1	225	Ć	0106	262	į	012F	303	Ŝ	015C	348
â	00E2	226	ć	0107	263	İ	0130	304	ŝ	015D	349
ã	00E3	227	Ĉ	0108	264	ı	0131	305	Ş	015E	350
ä	00E4	228	ĉ	0109	265	Ĵ	0134	308	ş	015F	351
å	00E5	229	Č	010C	268	ĵ	0135	309	Š	0160	352
æ	00E6	230	č	010D	269	Ķ	0136	310	š	0161	353
ç	00E7	231	Ď	010E	270	ķ	0137	311	Ţ	0162	354
è	00E8	232	ď	010F	271	Ĺ	0139	313	ţ	0163	355
é	00E9	233	Đ	0110	272	ĺ	013A	314	Ť	0164	356
ê	00EA	234	đ	0111	273	Ļ	013B	315	ť	0165	357
ë	00EB	235	Ē	0112	274	ļ	013C	316	Ũ	0168	360
ì	00EC	236	ē	0113	275	Ľ	013D	317	ũ	0169	361
í	00ED	237	Ĕ	0114	276	ľ	013E	318	Ū	016A	362
î	00EE	238	ĕ	0115	277	Ł	0141	321	ū	016B	363
ï	00EF	239	Ė	0116	278	ł	0142	322	Ŭ	016C	364
ð	00F0	240	ė	0117	279	Ń	0143	323	ŭ	016D	365
ñ	00F1	241	Ę	0118	280	ń	0144	324	Ů	016E	366

Glyph	Hex	Dec	Glyph	Hex	Dec	Glyph	Hex	Dec	Glyph	Hex	Dec
ů	016F	367	Ǣ	01E2	482	◌̮	032E	814	Ṙ	1E58	7768
Ű	0170	368	ǣ	01E3	483	◌̱	0331	817	ṙ	1E59	7769
ű	0171	369	Ǧ	01E6	486	π	03C0	960	Ṛ	1E5A	7770
Ų	0172	370	ǧ	01E7	487	Ḍ	1E0C	7692	ṛ	1E5B	7771
ų	0173	371	ǰ	01F0	496	ḍ	1E0D	7693	Ṝ	1E5C	7772
Ŷ	0176	374	Ș	0218	536	Ḏ	1E0E	7694	ṝ	1E5D	7773
ŷ	0177	375	ș	0219	537	ḏ	1E0F	7695	Ṟ	1E5E	7774
Ÿ	0178	376	ə	0259	601	Ḕ	1E14	7700	ṟ	1E5F	7775
Ź	0179	377	ɵ	0275	629	ḕ	1E15	7701	Ṣ	1E62	7778
ź	017A	378	ʰ	02B0	688	Ḗ	1E16	7702	ṣ	1E63	7779
Ż	017B	379	ʾ	02BE	702	ḗ	1E17	7703	Ṭ	1E6C	7788
ż	017C	380	ʿ	02BF	703	Ḥ	1E24	7716	ṭ	1E6D	7789
Ž	017D	381	ˆ	02C6	710	ḥ	1E25	7717	Ṯ	1E6E	7790
ž	017E	382	ˇ	02C7	711	Ḫ	1E2A	7722	ṯ	1E6F	7791
Ǝ	018E	398	˘	02D8	728	ḫ	1E2B	7723	Ẏ	1E8E	7822
ƒ	0192	402	˙	02D9	729	Ḵ	1E34	7732	ẏ	1E8F	7823
Ɵ	019F	415	˚	02DA	730	ḵ	1E35	7733	Ẓ	1E92	7826
Ơ	01A0	416	˛	02DB	731	Ḷ	1E36	7734	ẓ	1E93	7827
ơ	01A1	417	˜	02DC	732	ḷ	1E37	7735	Ẕ	1E94	7828
Ư	01AF	431	˝	02DD	733	Ḹ	1E38	7736	ẕ	1E95	7829
ư	01B0	432	◌̀	0300	768	ḹ	1E39	7737	ẖ	1E96	7830
Ǎ	01CD	461	◌́	0301	769	Ḻ	1E3A	7738	Ẽ	1EBC	7868
ǎ	01CE	462	◌̂	0302	770	ḻ	1E3B	7739	ẽ	1EBD	7869
Ǐ	01CF	463	◌̃	0303	771	Ṁ	1E40	7744	Ỳ	1EF2	7922
ǐ	01D0	464	◌̄	0304	772	ṁ	1E41	7745	ỳ	1EF3	7923
Ǒ	01D1	465	◌̆	0306	774	Ṃ	1E42	7746	Ỹ	1EF8	7928
ǒ	01D2	466	◌̇	0307	775	ṃ	1E43	7747	ỹ	1EF9	7929
Ǔ	01D3	467	◌̈	0308	776	Ṅ	1E44	7748	‍	200D	8205
ǔ	01D4	468	◌̊	030A	778	ṅ	1E45	7749	–	2013	8211
Ǖ	01D5	469	◌̋	030B	779	Ṇ	1E46	7750	—	2014	8212
ǖ	01D6	470	◌̌	030C	780	ṇ	1E47	7751	'	2018	8216
Ǘ	01D7	471	◌̐	0310	784	Ṉ	1E48	7752	'	2019	8217
ǘ	01D8	472	◌̣	0323	803	ṉ	1E49	7753	‚	201A	8218
Ǚ	01D9	473	◌̤	0324	804	Ṑ	1E50	7760	"	201C	8220
ǚ	01DA	474	◌̥	0325	805	ṑ	1E51	7761	"	201D	8221
Ǜ	01DB	475	◌̧	0327	807	Ṓ	1E52	7762	„	201E	8222
ǜ	01DC	476	◌̨	0328	808	ṓ	1E53	7763	†	2020	8224

Glyph	Hex	Dec	Glyph	Hex	Dec	Glyph	Hex	Dec	Glyph	Hex	Dec
‡	2021	8225	Ì̤	F10A	61706	Í̥	F12D	61741	Ṭ	F150	61776
•	2022	8226	ì̤	F10B	61707	L̥̀	F12E	61742	ṭ	F151	61777
…	2026	8230	Ĩ̤	F10C	61708	l̥̀	F12F	61743	M̆	F152	61778
‰	2030	8240	ĩ̤	F10D	61709	L̥	F130	61744	m̆	F153	61779
‹	2039	8249	Ĭ̤	F10E	61710	l̥	F131	61745	Ň	F154	61780
›	203A	8250	ĭ̤	F10F	61711	L̥̄	F132	61746	ň	F155	61781
⁄	2044	8260	Ú̥	F110	61712	Ị̥	F133	61747	Ř	F156	61782
€	20AC	8364	ú̥	F111	61713	Ĺ̥	F134	61748	ř	F157	61783
™	2122	8482	Ù̥	F112	61714	Í̥	F135	61749	Ĵ	F158	61784
Ω	2126	8486	ù̥	F113	61715	L̥̀	F136	61750	ĵ	F159	61785
∂	2202	8706	Ũ̥	F114	61716	l̥̀	F137	61751	Ĵ	F15A	61786
Δ	2206	8710	ũ̥	F115	61717	Ẽ̥	F138	61752	Z̩	F15B	61787
∏	220F	8719	Ŭ̥	F116	61718	ẽ̥	F139	61753	z̩	F15C	61788
∑	2211	8721	ŭ̥	F117	61719	Ĕ̥	F13A	61754	R̥̂	F15D	61789
−	2212	8722	Ŕ̥	F118	61720	ĕ̥	F13B	61755	r̥̂	F15E	61790
√	221A	8730	ŕ̥	F119	61721	Õ̥	F13C	61756	ᴴ	F15F	61791
∞	221E	8734	Ṛ̀	F11A	61722	õ̥	F13D	61757	ÃĨ	F160	61792
∫	222B	8747	ṛ̀	F11B	61723	Ŏ̥	F13E	61758	ãĩ	F161	61793
≈	2248	8776	Ŕ̥	F11C	61724	ŏ̥	F13F	61759	ÃŨ	F162	61794
≠	2260	8800	ŕ̥	F11D	61725	M̌	F140	61760	ãũ	F163	61795
≤	2264	8804	Ṛ̀	F11E	61726	m̌	F141	61761	˘	F1F8	61944
≥	2265	8805	ṛ̀	F11F	61727	L̃	F142	61762	–	F1F9	61945
◊	25CA	9674	R̥	F120	61728	Ĩ	F143	61763	ᷖ	F1FA	61946
			r̥	F121	61729	L̥̃	F144	61764	ᷗ	F1FB	61947
			R̥̄	F122	61730	Ĩ̥	F145	61765	ᷘ	F1FC	61948
Private Use Area:			r̥̄	F123	61731	K̲H̲	F146	61766	ᷙ	F1FD	61949
Ḁ́	F100	61696	Ŕ̥	F124	61732	K̲h̲	F147	61767	ᷚ	F1FE	61950
ḁ́	F101	61697	ŕ̥	F125	61733	k̲h̲	F148	61768	ᷛ	F1FF	61951
Ḁ̀	F102	61698	Ṛ̀	F126	61734	S̲	F149	61769	Adobe PUA:		
ḁ̀	F103	61699	ṛ̀	F127	61735	s̲	F14A	61770	ȷ	F6BE	63166
Ḁ̃	F104	61700	Ŕ̥	F128	61736	H̲	F14B	61771	،	F6C3	63171
ḁ̃	F105	61701	ŕ̥	F129	61737	Ḥ	F14C	61772	fi	FB01	64257
Ḁ̆	F106	61702	Ṛ̀	F12A	61738	ḥ	F14D	61773	fl	FB02	64258
ḁ̆	F107	61703	ṛ̀	F12B	61739	Ṣ	F14E	61774	Apple PUA:		
Í̥	F108	61704	Ĺ̥	F12C	61740	ṣ	F14F	61775		F8FF	63743
í̥	F109	61705									

1. The letters F15F (ᴴ) and F6BE (ȷ) are only accessible as Private Use Area codes.
2. The internal Private Use Area comma-accent F6C3 (،) should never be used.

Combining Diacritical Marks:	Stackable Diacritics: * = PUA code point		
		A◌̊ = Å	Đ = Đ (no GS!)
		a◌̊ = å	đ = đ (no GS!)
		A◌̨ = Ą	
◌̀ 0300	A◌̄ = Ā	a◌̨ = ą	E◌̄ = Ē
◌́ 0301	a◌̄ = ā	A◌̆=Ă	e◌̄ = ē
◌̂ 0302	A◌̆ = Ă	a◌̆=ă	E◌̆ = Ĕ
◌̃ 0303	a◌̆ = ă		e◌̆ = ĕ
◌̄ 0304	A◌̀ = À	AĬ = AĬ (no GS!)	E◌̀ = È
◌̆ 0306	a◌̀ = à	aĩ = aĩ (no GS!)	e◌̀ = è
◌̇ 0307	A◌́ = Á	AŨ = AŨ (no GS!)	E◌́ = É
◌̈ 0308	a◌́ = á	aũ = aũ (no GS!)	e◌́ = é
◌̊ 030A	A◌̃ = Ã		E◌̃ = Ẽ
◌̋ 030B	a◌̃ = ã	C◌̂ = Ĉ	e◌̃ = ẽ
◌̌ 030C	A◌̂ = Â	c◌̂ = ĉ	E◌̂ = Ê
◌̐ 0310	a◌̂ = â	C◌̨ = Ç	e◌̂ = ê
◌̣ 0323	A◌̈ = Ä	c◌̨ = ç	E◌̈ = Ë
◌̤ 0324	a◌̈ = ä	C◌́ = Ć	e◌̈ = ë
◌̥ 0325	A◌̄◌̆ = Ǟ *	c◌́ = ć	E◌̄◌̆ = Ḛ *
◌̧ 0327	a◌̄◌̆ = ǟ *	C◌̂ = Č	e◌̄◌̆ = ḛ *
◌̨ 0328	A◌̄◌̀ = Ȁ *	c◌̂ = č	E◌̄◌̀ = Ȅ
◌̮ 032E	a◌̄◌̀ = ȁ *		e◌̄◌̀ = ȅ
◌̱ 0331	A◌̄◌́ = Ȃ *	D◌̣ = Ḍ	E◌̄◌́ = Ȇ
	a◌̄◌́ = ȃ *	d◌̣ = ḍ	e◌̄◌́ = ȇ
	A◌̄◌̃ = Ȃ *	D◌̱ = Ḏ	E◌̄◌̃ = Ȇ *
	a◌̄◌̃ = ȃ *	d◌̱ = ḏ	e◌̄◌̃ = ȇ *
	Æ = Æ (no GS!)	D◌̆ = Ď (caron!)	Ǝ = Ǝ (no GS!)
	æ = æ (no GS!)	d◌̆ = ď (caron!)	ə = Ǝ (no GS!)
	Æ◌̄ = Ǣ	Đ = Đ (no GS!)	E◌̇ = Ė
	æ◌̄ = ǣ	ð = ð (no GS!)	e◌̇ = ė

Column 1:

E◌̦ = Ȩ
e◌̦ = ȩ
E◌̌ = Ĕ
e◌̌ = ĕ

G◌̇ = Ġ
g◌̇ = ġ
G◌̌ = Ğ
g◌̌ = ğ
G◌̌ = Ğ
g◌̌ = ğ
G◌̂ = Ĝ
g◌̂ = ĝ
G◌̦ = Ģ (cedilla!)
g◌̦ = ģ (cedilla!)

h◌̲ = h̲
H◌̲ = H̲ *
H◌̦ = Ḩ
h◌̦ = ẖ
H◌̦ = Ḩ
h◌̦ = ḥ
H◌̣ = Ḥ *
h◌̣ = ḥ *
Ĥ◌̂ = Ĥ
ĥ◌̂ = ĥ

ᴴ = ᴴ (no GS!) *
ʰ = ʰ (no GS!)

Column 2:

I◌̄ = Ī
i◌̄ = ī
I◌̌ = Ĭ
i◌̌ = ĭ
I◌̀ = Ì
i◌̀ = ì
I◌́ = Í
i◌́ = í
I◌̃ = Ĩ
i◌̃ = ĩ
I◌̂ = Î
i◌̂ = î
I◌̈ = Ï
i◌̈ = ï
I◌̄◌̌ = Ī̆ *
i◌̄◌̌ = ī̆ *
I◌̄◌̀ = Ī̀ *
i◌̄◌̀ = ī̀ *
I◌̄◌́ = Ī́ *
i◌̄◌́ = ī́ *
I◌̄◌̃ = Ī̃ *
i◌̄◌̃ = ī̃ *
I◌̨ = Į
i◌̨ = į
I◌̌ = Ǐ
i◌̌ = ǐ

İ = İ (no GS!)
ı = ı (no GS!)

Column 3:

J◌̌ = J̌ *
j◌̌ = ǰ
J◌́ = J́ *
j◌́ = j́ *
J◌̀ = J̀
j◌̀ = j̀
J◌̂ = Ĵ
j◌̂ = ĵ

ᴊ = ᴊ (no GS!)

K◌̲ = K̲
k◌̲ = k̲
K⁀h◌̲ = K̲h (ZWJ!) *
k⁀h◌̲ = k̲h (ZWJ!) *
K⁀H◌̲ = K̲H (ZWJ!) *
K◌̦ = Ķ (cedilla!)
k◌̦ = ķ (cedilla!)

L◌̲ = L̲
l◌̲ = l̲
L◌̦ = Ḷ
l◌̦ = ḷ
L◌̥ = L̥ *
l◌̥ = l̥ *
L◌̦◌̄ = Ḹ
l◌̦◌̄ = ḹ
L◌̥◌̄ = L̥̄ *
l◌̥◌̄ = l̥̄ *
L◌̦◌̀ = Ḷ̀ *
l◌̦◌̀ = ḷ̀ *
L◌̥◌̀ = L̥̀ *
l◌̥◌̀ = l̥̀ *

Column 4:

L◌̦◌́ = Ḷ́ *
l◌̦◌́ = ḷ́ *
L◌̥◌́ = Ĺ̥ *
l◌̥◌́ = ĺ̥ *
L◌̃ = L̃ *
l◌̃ = l̃ *
L◌̥ = L̥ *
l◌̥ = l̥ *
L◌́ = Ĺ
l◌́ = ĺ

L◌̦ = Ļ (cedilla!)
l◌̦ = ļ (cedilla!)
L◌̌ = Ľ (caron!)
l◌̌ = ľ (caron!)
Ł = Ł (no GS!)
ł = ł (no GS!)

M◌̇ = Ṁ
m◌̇ = ṁ
M◌̦ = Ṃ
m◌̦ = ṃ
M◌̌ = M̌ *
m◌̌ = m̌ *
M◌̌ = M̌ *
m◌̌ = m̌ *

N◌̲ = N̲
n◌̲ = n̲
N◌̌ = Ň *
n◌̌ = ň *

N˜• = Ñ	O᷈•́ = Ő	R•̥̀ = R̥̀ *	S• = S̲ *
n˜• = ñ	o᷈•́ = ő	r•̥̀ = r̥̀ *	s• = s̲ *
N•̇ = Ṅ	O᷈˜ = Õ *	R•̥̀ = R̥̀ *	Ŝ = Ŝ
n•̇ = ṅ	o᷈˜ = õ *	r•̥̀ = r̥̀ *	ŝ = ŝ
N• = Ṇ	O˝ = Ő	R•́ = Ŕ *	Š = Š
n• = ṇ	o˝ = ő	r•́ = ŕ *	š = š
Ń• = Ń	O�‌•̆=Ŏ	R•̥́ = Ŕ̥ *	
ń• = ń	o•̆=ŏ	r•̥́ = ŕ̥ *	
Ň• = Ň	Ơ = Ơ (no GS!)	R•̄̀ = R̲̀ *	S• = Ş (true cedilla!)
ň• = ň	ơ = ơ (no GS!)	r•̄̀ = r̲̀ *	s• = ş (true cedilla!)
N• = Ṇ (cedilla!)	Ø = Ø (no GS!)	R•̥̄̀ = R̥̲̀ *	Ş = Ş (no GS!)
n• = ṇ (cedilla!)	ø = ø (no GS!)	r•̥̄̀ = r̥̲̀ *	ş = ş (no GS!)
	Θ = Θ (no GS!)	R•̄́ = Ŕ̲ *	ß = ß (no GS!)
	θ = θ (no GS!)	r•̄́ = ŕ̲ *	
O᷈ = Ō	Œ = Œ (no GS!)	R•̥̄́ = Ŕ̥̲ *	T• = T̲
o᷈ = ō	œ = œ (no GS!)	r•̥̄́ = ŕ̥̲ *	t• = t̲
O•̆ = Ŏ		R•̂ = R̂ *	T• = Ṭ
o•̆ = ŏ	R• = R̲	r•̂ = r̂ *	t• = ṭ
O•̀ = Ò	r• = r̲	R•́ = Ŕ	T•̈ = Ṭ *
o•̀ = ò	R•̆ = Ř *	r•́ = ŕ	t•̈ = ṭ *
O•́ = Ó	r•̆ = ř *	R•̆ = Ř	
o•́ = ó	R•̇ = Ṙ	r•̆ = ř	T• = Ţ (cedilla!)
O˜• = Õ	r•̇ = ṙ	r•̆ = ř	t• = ţ (cedilla!)
o˜• = õ	R• = Ṛ	R• = Ṛ (cedilla!)	Ť•̆ = Ť (caron!)
Ô• = Ô	r• = ṛ	r• = ṛ (cedilla!)	t•̆ = ť (caron!)
ô• = ô	R•̥ = R̥ *		
O•̈ = Ö	r•̥ = r̥ *	S•́ = Ś	Þ = Þ (no GS!)
o•̈ = ö	R•̄ = R̄	s•́ = ś	þ = þ (no GS!)
O᷈•̆ = Ŏ *	r•̄ = r̄	S• = Ṣ	
o᷈•̆ = ŏ *	R•̥̄ = R̥̄ *	s• = ṣ	U᷈ = Ū
O᷈•̀ = Ò	r•̥̄ = r̥̄ *	S•̈ = Ṣ *	u᷈ = ū
o᷈•̀ = ò		s•̈ = ṣ *	

U◌̆ = Ŭ	U◌◌ = Ū̃ *	U̇ = U̇ (no GS!)	Z◌ = Ẕ
u◌̆ = ŭ	u◌◌ = ū̃ *	u̇ = u̇ (no GS!)	z◌ = ẕ
U◌ = Ù	U◌ = Ů		Z◌ = Ż
u◌ = ù	u◌ = ů		z◌ = ż
U◌ = Ú	U◌ = Ų	Y◌ = Ỹ	Z◌ = Ẓ
u◌ = ú	u◌ = ų	y◌ = ỹ	z◌ = ẓ
U◌ = Ũ	U◌ = Ű		Z◌◌ = Ẕ *
u◌ = ũ	u◌ = ű		z◌◌ = ẕ *
U◌ = Û	U◌◌ = Ŭ	Y◌ = Ẏ	Z◌ = Ž
u◌ = û	u◌◌ = ŭ	y◌ = ẏ	z◌ = ž
U◌ = Ü	U◌◌◌ = Ǘ	Y◌ = Ỳ	Z◌ = Ź
u◌ = ü	u◌◌◌ = ǘ	y◌ = ỳ	z◌ = ź
U◌◌ = Ŭ̀ *	U◌◌◌ = Ǜ	Y◌ = Ý	
u◌◌ = ŭ̀ *	u◌◌◌ = ǜ	y◌ = ý	
U◌◌ = Ù̆ *	U◌◌◌ = Ǚ	Y◌ = Ŷ	Adobe Ligatures:
u◌◌ = ù̆ *	u◌◌◌ = ǚ	y◌ = ŷ	fi = fi *
U◌◌ = Ű̆ *	U◌◌◌ = Ǖ	y◌ = ÿ	fl = fl *
u◌◌ = ű̆ *	u◌◌◌ = ǖ	Y◌ = Ÿ	

1. In order to trigger the ligature kh, you must insert the sign ↑ ("zero-width joiner" = ZWJ) between k and h: k↑ẖ = k̲h̲. The same applies for the ligatures Kh and KH. The ZWJ is not needed, when you encode k and h as separate, non-ligature letters: ḵẖ = k̲h̲.

2. The diacritics of modern languages with comma-accent (Latvian etc.) must be encoded with cedilla (not comma-accent) due to faulty prescriptions by the Unicode Consortium, and Şş is not encodable with cedilla at all, since Şş must be encoded with cedilla too.

4.5. List of Diacritics of Latin Alphabets covered by URW Palladio HOT

It is not intended to cover all possible diacritics of all languages based on the Latin script. But all diacritics required for <u>Pan-European</u> Windows Code Pages 1250, 1252, 1254, 1257 and for all Pan-European Macintosh Code Page are <u>fully</u> covered by URW Palladio HOT. However, with the help of the approx. 300 Indic diacritics, ISO 15919 and non-ISO, many other languages with Latin script are covered as well.

The following list is based on the handbook: DUDEN Satz und Korrektur, Mannheim 2003.

Afrikaans	usually written without diacritics
Albanian	Ç ç Ë ë
Azerbaijani	Ç ç (uppercase of ə not available, use Ǝ instead) ə Ğ ğ I ı İ i Ş ş Ü ü
Basque	not fully covered
Breton	C'h (not C'h) c'h (not c'h) È è É é Ê ê Ï ï Ñ ñ Ü ü
Catalan	À à Ç ç È è É é Í í Ï ï Ò ò Ó ó Ú ú Ü ü
Chinese (Pinyin)	Ü ü (for additional tone marks see page 3)
Croatian	Ć ć Č č Đ đ (Dj dj) Dž dž Š š Ž ž
Czech	Á á Č č Ď ď É é Ě ě Í í Ň ň Ó ó Ř ř Š š Ť ť Ú ú Ů ů Ý ý Ž ž
Danish	Å å Æ æ Ø ø (Ö ö)
Dutch	usually written without diacritics
English	usually written without diacritics
Esperanto	Ĉ ĉ Ĝ ĝ Ĥ ĥ Ĵ ĵ Ŝ ŝ Ŭ ŭ
Estonian	Ä ä Õ õ Ö ö Š š Ü ü Ž ž
Faeroese	Á á Æ æ Ð ð Í í Ó ó Ø ø Ú ú Ý ý
Finnish	Ä ä Ö ö Š š Ž ž
French	À à Â â Æ æ Ç ç È è É é Ê ê Ë ë Î î Ï ï Ô ô Œ œ Û û Ü ü
Frisian	Â â É é Ê ê ë ï Ô ô ö Ú ú Û û
Gaelic	À à É é Ì ì Ò ò Ó ó Ù ù
Galician	Á á É é Í í Ñ ñ Ó ó Ú ú Ü ü
German	Ä ä Ö ö Ü ü ß
Hausa	usually written without diacritics
Hungarian	Á á É é Í í Ó ó Ö ö Ő ő Ú ú Ü ü Ű ű
Icelandic	Á á Æ æ Ð ð É é Í í Ó ó Ö ö Þ þ Ú ú Ý ý
Indonesian	usually written without diacritics

Irish	Á á É é Í í Ó ó Ú ú
Italian	À à (Á á) È è É é Ì ì (Í í) Î î Ò ò Ó ó Ù ù (Ú ú)
Lappish	usually written without diacritics
Latin	Ā ā Ă ă Ǎ ǎ Ē ē Ĕ ĕ Ě ě Ī ī Ĭ ĭ Ǐ ǐ Ō ō Ŏ ŏ Ǒ ǒ Ū ū Ŭ ŭ Ǔ ǔ
Latvian	Ā ā Č č Ē ē Ģ ģ Ī ī Ķ ķ Ļ ļ Ņ ņ Ŗ ŗ Š š Ū ū (Ų ų) Ž ž
Lithuanian	Ą ą Č č Ė ė Ę ę Ļ ļ Š š Ū ū Ų ų Ž ž
Luxembourgian	Ä ä Ç ç É é Ë ë Ö ö Ü ü
Malagasy	usually written without diacritics
Maltese	not fully covered
Norwegian	À à Æ æ Å å È è É é Ò ò Ô ô Ø ø
Polish	Ą ą Ć ć Ę ę Ł ł Ń ń Ó ó Ś ś Ź ź Ż ż
Portuguese	À à Á á Â â Ã ã Ç ç (È è) É é (Ê ê) Í í (Î î Ï ï) (Ò ò) Ó ó Ô ô Õ õ (Ù ù) Ú ú
Rhaeto-Romanic	À à (Á á) Â â È è É é Ê ê Î î Ï ï Ò ò Ô ô Ö ö Ù ù Ü ü
Romanian	Â â Ă ă Î î Ș ș (not Ş ş) Ț ț
Samoan	not fully covered
Slovak	Á á Ä ä Č č Ď ď É é Í í Ĺ ĺ Ľ ľ Ň ň Ó ó Ô ô Ŕ ŕ Š š Ť ť Ú ú Ý ý Ž ž
Slovenian	Č č Š š Ž ž
Somali	usually written without diacritics
Sorbian (N.+O.)	Ć ć Č č Dź dź Ě ě Ł ł Ń ń Ó ó Ŕ ŕ Ř ř Ś ś Š š Ź ź Ž ž
Spanish	Á á É é Í í Ñ ñ Ó ó Ú ú Ü ü
Swahili	usually written without diacritics
Swedish	Ä ä Å å Ö ö
Tagalog	usually written without diacritics
Turkish	Â â Ç ç Ğ ğ I ı İ i Î î Ö ö Ș ș (not Ş ş) Û û Ü ü
Vietnamese	Â â Ă ă Đ đ Ê ê Ô ô Ơ ơ Ư ư (tone marks not covered)
Welsh	not fully covered

Palladio HOT covers the four <u>Windows</u> Pan-European Code Pages 1250, 1252, 1254, 1257 and the six <u>Macintosh</u> Code Pages ("US Roman", "Central Europe", "Croatian", "Icelandic", "Romanian", and "Turkish").

URW Palladio HOT Roman – Keyboard Layout for Code Page 1252

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
032		!	"	#	$	%	&	'	(	)	*	+	,	-	.	/
048	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
064	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
080	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
096	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
112	p	q	r	s	t	u	v	w	x	y	z	{	|	}	~	
128	€		‚	ƒ	„	…	†	‡	^	‰	Š	‹	Œ		Ž	
144		'	'	"	"	•	–	—	~	™	š	›	œ		ž	Ÿ
160		¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯
176	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
192	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
208	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
224	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
240	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

Notes:

1. The grey 7-bit ASCII range 032–127 (hex 20–7F) is the same for every Code Page, i.e. all the Code Pages differ only by the 8-bit ANSI range 128-255 (hex 80–FF).

2. The above chart shows the <u>key</u> codes. Example: For Œ, you key in the decimal code 140 (ALT 0140), whereas the internal code for Œ is hex 0152 (decimal 338).

Windows Code Page 1252 – Western Europe

128	€		,	ƒ	„	…	†	‡	^	‰	Š	‹	Œ		Ž	
144		'	'	"	"	•	–	—	~	TM	š	›	œ		ž	Ÿ
160		¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯
176	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
192	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
208	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
224	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
240	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

Windows Code Page 1250 – Eastern Europe

128	€		,		„	…	†	‡		‰	Š	‹	Ś	Ť	Ž	Ź
144		'	'	"	"	•	–	—		TM	š	›	ś	ť	ž	ź
160		ˇ	˘	Ł	¤	Ą	¦	§	¨	©	Ş	«	¬	-	®	Ż
176	°	±	˛	ł	´	µ	¶	·	¸	ą	ş	»	Ľ	˝	ľ	ż
192	Ŕ	Á	Â	Ă	Ä	Ĺ	Ć	Ç	Č	É	Ę	Ë	Ě	Í	Î	Ď
208	Đ	Ń	Ň	Ó	Ô	Ő	Ö	×	Ř	Ů	Ú	Ű	Ü	Ý	Ţ	ß
224	ŕ	á	â	ă	ä	Í	ć	ç	č	é	ę	ë	ě	í	î	ď
240	đ	ń	ň	ó	ô	ő	ö	÷	ř	ů	ú	ű	ü	ý	ţ	˙

Windows Code Page 1254 – Turkish

128	€		‚	ƒ	„	…	†	‡	ˆ	‰	Š	‹	Œ			
144		'	'	"	"	•	–	—	˜	™	š	›	œ			Ÿ
160		¡	¢	£	¤	¥	¦	§	¨	©	ª	«	¬	-	®	¯
176	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
192	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
208	Ğ	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	İ	Ş	ß
224	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
240	ğ	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ı	ş	ÿ

Windows Code Page 1257 – Baltic

128	€		‚		„	…	†	‡		‰		‹		¨	ˇ	¸
144		'	'	"	"	•	–	—		™		›		¯		˛
160			¢	£	¤		¦	§	Ø	©	Ŗ	«	¬	-	®	Æ
176	°	±	²	³	´	µ	¶	·	ø	¹	ŗ	»	¼	½	¾	æ
192	Ą	Į	Ā	Ć	Ä	Å	Ę	Ē	Č	É	Ź	Ė	Ģ	Ķ	Ī	Ļ
208	Š	Ń	Ņ	Ó	Ō	Õ	Ö	×	Ų	Ł	Ś	Ū	Ü	Ż	Ž	ß
224	ą	į	ā	ć	ä	å	ę	ē	č	é	ź	ė	ģ	ķ	ī	ļ
240	š	ń	ņ	ó	ō	õ	ö	÷	ų	ł	ś	ū	ü	ż	ž	˙

Macintosh Code Page "US Roman"

128	Ä	Å	Ç	É	Ñ	Ö	Ü	á	à	â	ä	ã	å	ç	é	è
144	ê	ë	í	ì	î	ï	ñ	ó	ò	ô	ö	õ	ú	ù	û	ü
160	†	°	¢	£	§	•	¶	ß	®	©	™	´	¨	≠	Æ	Ø
176	∞	±	≤	≥	¥	µ	∂	∑	∏	π	∫	ª	º	Ω	æ	ø
192	¿	¡	¬	√	ƒ	≈	Δ	«	»	…		À	Ã	Õ	Œ	œ
208	–	—	"	"	'	'	÷	◊	ÿ	Ÿ	/	€	‹	›	fi	fl
224	‡	·	‚	„	‰	Â	Ê	Á	Ë	È	Í	Î	Ï	Ì	Ó	Ô
240		Ò	Ú	Û	Ù	ı	ˆ	˜	¯	˘	˙	˚	¸	˝	˛	ˇ

Macintosh Code Page "Central Europe"

128	Ä	Ā	ā	É	Ą	Ö	Ü	á	ą	Č	ä	č	Ć	ć	é	Ź
144	ź	Ď	í	ď	Ē	ē	Ė	ó	ė	ô	ö	õ	ú	Ě	ě	ü
160	†	°	Ę	£	§	•	¶	ß	®	©	™	ę	¨	≠	ǵ	Į
176	į	Ī	≤	≥	ī	Ķ	∂	∑	ł	Ļ	ļ	Ľ	ľ	Ĺ	Í	Ņ
192	ņ	Ń	¬	√	ń	Ň	Δ	«	»	…		ň	Ő	Õ	ő	Ō
208	–	—	"	"	'	'	÷	◊	ō	Ŕ	ŕ	Ř	‹	›	ř	Ŗ
224	ŗ	Š	‚	„	š	Ś	ś	Á	Ť	ŧ	Í	Ž	ž	Ū	Ó	Ô
240	ū	Ů	Ú	ů	Ű	ű	Ų	ų	Ý	ý	ķ	Ż	Ł	ż	Ģ	ˇ

<u>Note:</u> Four additional Mac Code Pages ("Croatian", "Icelandic", "Romanian", and "Turkish")
– also fully covered by URW Palladio HOT – are not listed for lack of space.

5.1. URW Palladio HOT Roman

! " # $ % & ' ( ) * + , - . / 0 1 2 3 4 5 6 7 8 9

: ; < = > ? @ A B C D E F G H I J K L M N O

P Q R S T U V W X Y Z [\ ] ^ _ ` a b c d e f

g h i j k l m n o p q r s t u v w x y z { | } ~

¡ ¢ £ ¤ ¥ ¦ § ¨ © ª « ¬ - ® ¯ ° ± ² ³ ´ µ ¶ · ¸ ¹ º

» ¼ ½ ¾ ¿ À Á Â Ã Ä Å Æ Ç È É Ê Ë Ì Í Î Ï

Ð Ñ Ò Ó Ô Õ Ö × Ø Ù Ú Û Ü Ý Þ ß à á â ã

ä å æ ç è é ê ë ì í î ï ð ñ ò ó ô õ ö ÷ ø ù ú û

ü ý þ ÿ Ā ā Ă ă Ą ą Ć ć Ĉ ĉ Č č Ď ď Đ đ Ē ē

Ĕ ĕ Ė ė Ę ę Ě ě Ĝ ĝ Ğ ğ Ġ ġ Ģ ģ Ĥ ĥ Ĩ ĩ Ī ī Ĭ

ĭ Į į İ ı Ĵ ĵ Ķ ķ Ĺ ĺ Ļ ļ Ľ ľ Ł ł Ń ń Ņ ņ Ň ň Ō

ō Ŏ ŏ Ő ő Œ œ Ŕ ŕ Ŗ ŗ Ř ř Ś ś Ŝ ŝ Ş ş Š š Ţ ţ

Ť ť Ũ ũ Ū ū Ŭ ŭ Ů ů Ű ű Ų ų Ŷ ŷ Ÿ Ź ź Ż ż

Ž ž Ǝ ƒ Ɵ Ơ ơ Ư ư Ǎ ǎ Ǐ ǐ Ǒ ǒ Ǔ ǔ Ǖ ǖ Ǘ ǘ

Ǚ ǚ Ǜ ǜ Ǣ ǣ Ǧ ǧ ǰ Ş ş ə ɵ ʰ ʾ ʿ ˆ ˇ ˘ ˙ ˚ ˜ ˝ ˳

́ ̂ ̃ ̄ ̆ ̈ ̊ ̋ ̌ ̇ ̣ ̈ ̥ ̧ ̨ ̌ ̩ π Ḍ ḍ Ḏ

ḏ Ḕ ḕ Ḗ ḗ Ḥ ḥ Ḫ ḫ Ḵ ḵ Ḷ ḷ Ḹ ḹ Ḻ ḻ Ṁ ṁ Ṃ ṃ

Ṅ ṅ Ṇ ṇ Ṉ ṉ Ṍ ṍ Ṏ ṏ Ṙ ṙ Ṛ ṛ Ṝ ṝ Ṟ ṟ Ṣ ṣ Ṭ ṭ

Ṯ ṯ Ẏ ẏ Ẕ ẕ Ẑ ẑ ẖ Ẽ ẽ Ỳ ỳ Ỹ ỹ ⁀ – — ' ' ‚ " " „

† ‡ • … ‰ ‹ › ⁄ € ™ Ω ∂ Δ ∏ ∑ − √ ∞ ∫ ≈ ≠ ≤

≥ ◊ Ǻ ǻ Ầ ầ Ẫ ẫ Ẩ ẩ Ị ị Ĩ ĩ Ĩ ĩ Ỉ ỉ Ṹ ṹ Ǜ ǜ Ũ ũ

Ǔ ǔ Ṝ ṝ Ṛ ṛ Ṝ ṝ Ṛ ṛ Ṟ ṟ Ṝ ṝ Ṝ ṝ Ṝ ṝ Ṝ ṝ Ṝ ṝ Ḹ

ḹ Ḹ ḹ Ḻ ḻ Ḹ ḹ Ḹ ḹ Ẽ ẽ Ẽ ẽ Ȭ õ Ȭ õ Ṁ ṁ Ẽ ẽ

Ḻ ḻ ₭Ｈ ₭ｈ ｋｈ Ṣ ṣ Ḥ Ḥ ḥ Ṣ ṣ Ṭ ṭ Ṁ ṁ Ň ň Ř ř

Ĵ ĵ Ĵ Ẕ ẕ Ṝ ṝ ʰ Ａĩ ãĩ Ａũ ãũ ˘ ‾ ̭ ̮ ̯ ̰ ́ ̀ ˌ ,

ﬁ ﬂ

! " # $ % & ' ( ) * + , - . / 0 1 2 3 4 5 6 7 8 9

: ; < = > ? @ A B C D E F G H I J K L M N O

P Q R S T U V W X Y Z [\ ] ^ _ ` a b c d e f

g h i j k l m n o p q r s t u v w x y z { | } ~   ¡

¢ £ ¤ ¥ ¦ § ¨ © ª « ¬ - ® ¯ ° ± ² ³ ´ µ ¶ · ¸ ¹ º »

¼ ½ ¾ ¿ À Á Â Ã Ä Å Æ Ç È É Ê Ë Ì Í Î Ï Ð

Ñ Ò Ó Ô Õ Ö × Ø Ù Ú Û Ü Ý Þ ß à á â ã ä å

æ ç è é ê ë ì í î ï ð ñ ò ó ô õ ö ÷ ø ù ú û ü ý þ ÿ

Ā ā Ă ă Ą ą Ć ć Ĉ ĉ Č č Ď ď Đ đ Ē ē Ĕ ĕ Ė ė Ę

ę Ě ě Ĝ ĝ Ğ ğ Ġ ġ Ģ ģ Ĥ ĥ Ĩ ĩ Ī ī Ĭ ĭ Į į İ ı Ĵ ĵ

Ķ ķ Ĺ ĺ Ļ ļ Ľ ľ Ł ł Ń ń Ņ ņ Ň ň Ō ō Ŏ ŏ Ő ő Œ

œ Ŕ ŕ Ŗ ŗ Ř ř Ś ś Ŝ ŝ Ş ş Š š Ţ ţ Ť ť Ũ ũ Ū ū

Ŭ ŭ Ů ů Ű ű Ų ų Ŷ ŷ Ÿ Ź ź Ż ż Ž ž Ǝ ƒ Ө Ơ ơ

Ư ư Ǎ ǎ Ǐ ǐ Ǒ ǒ Ǔ ǔ Ǖ ǖ Ǘ ǘ Ǚ ǚ Ǜ ǜ Ǽ ǽ Ǧ ǧ

ǰ Ș ș ə θ ʰ ˒ ˓ ˆ ˇ ˘ ˙ ˚ ˛ ˜ ˝ ̀ ́ ̂ ̃ ̄ ̆ ̇ ̈ ̊ ̋

̌ ̍ ̣ ̤ ̥ ̦ ̧ ̨ ̩ ▄ π Ḍ ḍ Ḏ ḏ Ḕ ḕ Ḗ ḗ Ḫ ḫ Ḥ ḥ

Ḵ ḵ Ḷ ḷ Ḹ ḹ Ḻ ḻ Ṁ ṁ Ṃ ṃ Ṅ ṅ Ṇ ṇ Ṉ ṉ Ṑ ṑ Ṓ ṓ

Ṙ ṙ Ṛ ṛ Ṝ ṝ Ṟ ṟ Ṣ ṣ Ṭ ṭ Ṯ ṯ Ẏ ẏ Ẓ ẓ Ẕ ẕ ẖ Ẽ ẽ Ỳ

ỳ Ỹ ỹ ⁺ – — ' ' ‚ " " „ † ‡ • … ‰ ‹ › ⁄ € ™ Ω

∂ Δ ∏ Σ − √ ∞ ∫ ≈ ≠ ≤ ≥ ◊ Á́ á́ À̀ à̀ Ã̃ ã̃ Ă̆ ă̆ Í́ í́

Ì̀ ì̀ Ĩ̃ ĩ̃ Ĭ̆ ĭ̆ Ú́ ú́ Ù̀ ù̀ Ũ̃ ũ̃ Ŭ̆ ŭ̆ Ŕ́ ŕ́ Ŕ̀ ŕ̀ Ŕ́ ŕ́ Ŕ̀ ŕ̀ Ŗ ŗ

Ṝ ṝ Ŕ̊ ŕ̊ R̀̊ r̀̊ Ŕ̂̊ ŕ̂̊ R̃̊ r̃̊ Ĺ ĺ Ĺ̀ ĺ̀ Ļ ļ Ḹ ḹ Ĺ́ ĺ́ Ĺ̀ ĺ̀ Ẽ̄ ẽ̄ Ĕ̄

ĕ̄ Ȭ ȭ Ŏ̄ ŏ̄ Ṁ̆ ṁ̆ L̃ l̃ Ľ̃ ľ̃ KH Kh kh S s H Ḥ ḥ Ṣ ṣ

Ṯ ṯ Ṁ̆ ṁ̆ Ň ň Ř ř J́ j́ J̌ ǰ Z̲ z̲ Ŗ̊ ŗ̊ ʰ ÃĪ ãī ÃŨ ãũ ˘ ˗

̪ ˘ ⁻ ̲ ˊ ˋ ɟ , fi fl

5.3. URW Palladio HOT Bold

! " # $ % & ' ( ) * + , - . / 0 1 2 3 4 5 6 7 8 9

: ; < = > ? @ A B C D E F G H I J K L M N

O P Q R S T U V W X Y Z [\ ] ^ _ ` a b c d

e f g h i j k l m n o p q r s t u v w x y z { |

} ~ ¡ ¢ £ ¤ ¥ ¦ § ¨ © ª « ¬ - ® ¯ ° ± ² ³ ´ µ ¶

· ¸ ¹ º » ¼ ½ ¾ ¿ À Á Â Ã Ä Å Æ Ç È É Ê Ë Ì

Í Î Ï Ð Ñ Ò Ó Ô Õ Ö × Ø Ù Ú Û Ü Ý Þ ß à

á â ã ä å æ ç è é ê ë ì í î ï ð ñ ò ó ô õ ö ÷ ø

ù ú û ü ý þ ÿ Ā ā Ă ă Ą ą Ć ć Ĉ ĉ Č č Ď ď Đ

đ Ē ē Ĕ ĕ Ė ė Ę ę Ě ě Ĝ ĝ Ğ ğ Ġ ġ Ģ ģ Ĥ ĥ Ĩ

ĩ Ī ī Ĭ ĭ Į į İ ı Ĵ ĵ Ķ ķ Ĺ ĺ Ļ ļ Ľ ľ Ł ł Ń ń Ņ ņ

Ň ň Ō ō Ŏ ŏ Ő ő Œ œ Ŕ ŕ Ŗ ŗ Ř ř Ś ś Ŝ ŝ Ş

ş Š š Ţ ţ Ť ť Ũ ũ Ū ū Ŭ ŭ Ů ů Ű ű Ų ų Ŷ ŷ Ÿ

Ź ź Ż ż Ž ž Ǝ ƒ Ɵ Ơ ơ Ư ư Ǎ ǎ Ǐ ǐ Ǒ ǒ Ǔ ǔ Ǖ

ǖ Ǘ ǘ Ǚ ǚ Ǜ ǜ Ǽ ǽ Ǧ ǧ ǰ Ș ș ə ɵ ʰ ˀ ˁ ˆ ˇ ˘

˚ ˛ ˜ ˝ ̀ ́ ̂ ̃ ̄ ̅ ̆ ̇ ̈ ̊ ̋ ̌ ̌ ̣ ̤ ̥ ̦ ̧ ̨ ̱

π Ḍ ḍ Ḏ ḏ Ḕ ḕ Ḗ ḗ Ḥ ḥ Ḫ ḫ Ḵ ḵ Ḷ ḷ Ḹ ḹ Ḻ ḻ

Ṁ ṁ Ṃ ṃ Ṅ ṅ Ṇ ṇ Ṉ ṉ Ṍ ṍ Ṓ ṓ Ṙ ṙ Ṛ ṛ Ṝ ṝ

Ṟ ṟ Ṣ ṣ Ṭ ṭ Ṯ ṯ Ẏ ẏ Ẓ ẓ Ẕ ẕ ẖ Ẽ ẽ Ỳ ỳ Ỹ ỹ † –

— ' ' ‚ " " „ † ‡ • … ‰ ‹ › / € ™ Ω ∂ Δ ∏ ∑

− √ ∞ ∫ ≈ ≠ ≤ ≥ ◊ Á á À à Ã ã Ǎ ǎ Í í Ì ì Ĩ ĩ Ǐ

ǐ Ú ú Ù ù Ũ ũ Ǔ ǔ Ŕ ŕ Ṛ ṛ Ŕ ŕ Ṛ ṛ Ŗ ŗ Ṝ ṝ Ŕ

ŕ Ṛ ṛ Ŕ ŕ Ṛ ṛ Ĺ ĺ Ḷ ḷ Ļ ļ Ḹ ḹ Ĺ ĺ Ḷ ḷ Ḙ ḙ Ḙ ḙ Ö

õ Ǒ ǒ Ṁ ṁ Ḽ ḽ Ḽ ḽ K̲H̲ K̲h̲ k̲h̲ S̲ s̲ H̱ Ḥ ḥ Ṣ ṣ Ṭ

ṭ M̌ m̌ Ň ň Ř ř Ĵ ĵ J̌ Z̲ z̲ R̥ r̥ ʜ A͠Ĭ a͠ĭ A͡Ū a͡ū ◡ −

◡ ◡ ◡ ◡ ◡ ◡ ◡ ◡ , fi fl

5.4. URW Palladio HOT Bold Italic

! ″ # $ % & ′ ( ) * + , - . / 0 1 2 3 4 5 6 7 8 9

: ; < = > ? @ A B C D E F G H I J K L M N O

P Q R S T U V W X Y Z [\ ] ^ _ ` a b c d e

f g h i j k l m n o p q r s t u v w x y z { | } ~

¡ ¢ £ ¤ ¥ ¦ § ¨ © ª « ¬ - ® ¯ ° ± ² ³ ´ µ ¶ · ¸ ¹

º » ¼ ½ ¾ ¿ À Á Â Ã Ä Å Æ Ç È É Ê Ë Ì Í Î Ï

Ð Ñ Ò Ó Ô Õ Ö × Ø Ù Ú Û Ü Ý Þ ß à á â ã

ä å æ ç è é ê ë ì í î ï ð ñ ò ó ô õ ö ÷ ø ù ú û

ü ý þ ÿ Ā ā Ă ă Ą ą Ć ć Ĉ ĉ Č č Ď ď Đ đ Ē ē

Ĕ ĕ Ė ė Ę ę Ě ě Ĝ ĝ Ğ ğ Ġ ġ Ģ ģ Ĥ ĥ Ĩ ĩ Ī ī Ĭ ĭ

Į į İ ı Ĵ ĵ Ķ ķ Ĺ ĺ Ļ ļ Ľ ľ Ł ł Ń ń Ņ ņ Ň ň Ō ō

Ŏ ŏ Ő ő Œ œ Ŕ ŕ Ŗ ŗ Ř ř Ś ś Ŝ ŝ Ş ş Š š Ţ ţ

Ť ť Ũ ũ Ū ū Ŭ ŭ Ů ů Ű ű Ų ų Ŷ ŷ Ÿ Ź ź Ż ż Ž

ž Ǝ ƒ Θ Ơ ơ Ư ư Ǎ ǎ Ǐ ǐ Ǒ ǒ Ǔ ǔ Ǖ ǖ Ǘ ǘ Ǚ

ǚ Ǜ ǜ Ǣ ǣ Ǧ ǧ ǰ Ş ş ə ɵ ʰ ˀ ˁ ˆ ˇ ˘ ˙ ˚ ˛ ˜ ˝ ̣ ́

̂ ̃ ̄ ̆ ̇ ̈ ̊ ̋ ̌ ̉ ̣ ̤ ̦ ̧ ̨ ̩ ̱ π Ḍ ḍ Ḏ ḏ

Ḕ ḕ Ḗ ḗ Ḥ ḥ Ḫ ḫ Ḵ ḵ Ḷ ḷ Ḹ ḹ Ḻ ḻ Ṁ ṁ Ṃ ṃ Ṅ ṅ

Ṇ ṇ Ṉ ṉ Ṑ ṑ Ṓ ṓ Ṙ ṙ Ṛ ṛ Ṝ ṝ Ṟ ṟ Ṣ ṣ Ṭ ṭ Ṯ ṯ

Ẏ ẏ Ẓ ẓ Ẕ ẕ ẖ Ẽ ẽ Ỳ ỳ Ỹ ỹ † – — ' ' ‚ " " „ † ‡

• … ‰ ‹ › ⁄ € ™ Ω ∂ Δ ∏ ∑ − √ ∞ ∫ ≈ ≠ ≤ ≥ ◊

Ǻ ǻ Ầ ầ Ẫ ẫ Ẵ ẵ Ấ ấ Ì ì Ĩ ĩ Ǐ ǐ Ű ű Ù ù Ữ ữ Ŭ

ŭ Ŕ ŕ Ṟ ṟ Ŕ́ ŕ́ Ṙ̀ ṙ̀ R r Ṝ ṝ Ŕ ŕ Ṙ ṙ Ŕ̂ ŕ̂ Ṙ̀ ṙ̀ Ḹ ḹ

Ḷ ḷ Ḽ ḽ Ḹ ḹ Ĺ ĺ Ḷ̀ ḷ̀ Ễ ễ Ḝ ḝ Ỗ ỗ Ỗ ỗ Ṁ̌ ṁ̌ Ḽ ḽ Ḽ

ḽ Ḵ̱ Ḵ̱ ḵ̱ Ṣ̱ ṣ̱ Ḥ Ḥ ḥ Ṣ̣ ṣ̣ Ṭ̣ ṭ̣ Ṁ̆ ṁ̆ Ṅ̆ ṅ̆ Ř ř J́ j́

J̌ Ẕ ẕ Ṟ̣ ṟ̣ ʜ ĀĪ āī ĀŪ āū ˘ ˗ ̮ ̯ ̱ ̲ ̭ ̬ ɟ , fi

fl

! " # $ % & ' ( ) * + , - . / 0 1 2 3 4 5 6 7 8 9

: ; < = > ? @ A B C D E F G H I J K L M N O

P Q R S T U V W X Y Z [\ ] ^ _ ` A B C D E F

G H I J K L M N O P Q R S T U V W X Y Z { | } ~   ¡

¢ £ ¤ ¥ ¦ § ¨ © ª « ¬ - ® ¯ ° ± ² ³ ´ µ ¶ · ‚ ¹ º

» ¼ ½ ¾ ¿ À Á Â Ã Ä Å Æ Ç È É Ê Ë Ì Í Î Ï

Ð Ñ Ò Ó Ô Õ Ö × Ø Ù Ú Û Ü Ý Þ SS À Á Â Ã

Ä Å Æ Ç È É Ê Ë Ì Í Î Ï Ð Ñ Ò Ó Ô Õ Ö ÷ Ø Ù Ú Û Ü

Ý Þ Ÿ Ā ā Ă ă Ą ą Ć ć Ĉ ĉ Č č Ď ď Đ đ Ē ē Ĕ

ĕ Ė ė Ę ę Ě ě Ĝ ĝ Ğ ğ Ġ ġ Ģ ģ Ĥ ĥ Ĩ ĩ Ī ī Ĭ ĭ Į į

İ ı Ĵ ĵ Ķ ķ Ĺ ĺ Ļ ļ Ľ ľ Ł ł Ń ń Ņ ņ Ň ň Ō ō Ŏ

ŏ Ő ő Œ œ Ŕ ŕ Ŗ ŗ Ř ř Ś ś Ŝ ŝ Ş ş Š š Ţ ţ Ť ť

Ũ ũ Ū ū Ŭ ŭ Ů ů Ű ű Ų ų Ŷ ŷ Ÿ Ź ź Ż ż Ž ž Ǝ

ƒ Ɵ Ơ ơ Ư ư Ǎ ǎ Ǐ ǐ Ǒ ǒ Ǔ ǔ Ǖ ǖ Ǘ ǘ Ǚ ǚ Ǜ ǜ

Ǣ ǣ Ǧ ǧ ǰ Ș ș Ɜ θ ʜ ʾ ʿ ˆ ˇ ˘ ˙ ˚ ˛ ˜ ˝ ̀ ́ ̂ ̃ ̄

̆ ̇ ̈ ̊ ̋ ̌ ̑ ̣ ̤ ̦ ̧ ̨ ̩ π Ḍ ḍ Ḏ ḏ Ḕ ḕ Ḗ

ḗ Ḥ ḥ Ḧ ḧ Ḵ ḵ Ḷ ḷ Ḹ ḹ Ḻ ḻ Ḿ ḿ Ṁ ṁ Ṅ ṅ Ṇ ṇ

Ṉ ṉ Ṍ ṍ Ṏ ṏ Ṙ ṙ Ṛ ṛ Ṝ ṝ Ṟ ṟ Ṣ ṣ Ṭ ṭ Ṱ ṱ Ẏ ẏ Ẓ

ẓ Ẕ ẕ ẖ Ẽ ẽ Ỳ ỳ Ỹ ỹ ʼ – — ' ' ‚ " " „ † ‡ • …

‰ ‹ › ⁄ € ™ Ω ∂ Δ ∏ Σ − √ ∞ ∫ ≈ ≠ ≤ ≥ ◊ Á́ á́

Ā̀ ā̀ Ã̄ ã̄ Ă̄ ă̄ Í̄ í̄ Ì̄ ì̄ Ĩ̄ ĩ̄ Ǐ̄ ǐ̄ Ú́ ú́ Ù̄ ù̄ Ũ̄ ũ̄ Ǔ̄ ǔ̄ Ŕ̩ ŕ̩

Ř̩ ř̩ Ŕ̩́ ŕ̩́ R̩̀ r̩̀ Ṛ̨ ṛ̨ R̩̄ r̩̄ Ŕ̥́ ŕ̥́ R̥̀ r̥̀ Ĺ̩ ĺ̩ L̩̀ l̩̀ L̩

l̩ L̩̄ l̩̄ Ĺ̥ ĺ̥ L̥̀ l̥̀ Ẽ̄ ẽ̄ Ĕ̄ ĕ̄ Ȭ ȭ Ŏ̄ ŏ̄ M̌ m̌ L̃ l̃ Ľ̩ ľ̩ KH

Kh KH Ṣ́ ṣ́ Ḥ̌ ḥ̌ Ṣ̌ ṣ̌ Ṭ̣ ṭ̣ M̆ m̆ Ň̩ ň̩ Ř̩ ř̩ J́ j́ J̌ z̩ Z̩ z̩

R̥̂ r̥̂ ᴴ ĀI̯ āi̯ ĀU̯ āu̯ ̯ ̠ ̤̄ ̌ ̄̌ ̥̄ x̱ x̀ ᴊ ‚ fi fl

6. Quick-and-Dirty Savvy-Test for URW Palladio HOT

1.1   *Try to enter ṝ via character map, <u>or</u>*

1.2   *try to enter U+1E5D via keyboard.*

1.3   *If you <u>see</u> ṝ with <u>dot</u> below, your word processor is <u>partially</u> Unicode-savvy.*

2.1   *Try to enter ṟ́ via character map, <u>or</u>*

2.2   *try to enter U+F129 via keyboard.*

2.3   *If you <u>see</u> ṟ́ with <u>ring</u> below, your word processor is <u>fully</u> Unicode-savvy.*

3.1   *Try to enter r ● ‾● via character map, <u>or</u>*

3.2   *try to enter U+0072, U+0323, U+0304 via keyboard.*

3.3   *If you <u>see</u> ṝ with <u>dot</u> below, your word processor is <u>partially</u> OpenType-savvy, i.e. it is Stacking-savvy.*

4.1   *Try to enter r ● ‾● ́● via character map, <u>or</u>*

4.2   *try to enter U+0072, U+0325, U+0304, U+0301 via keyboard.*

4.3   *If you <u>see</u> ṟ́ with <u>ring</u> below, your word processor is <u>fully</u> OpenType-savvy.*

This savvy-test is based on the fact that Unicode and PUA glyphs are not handled consistently by all word processors.

7. OpenType Horror Picture Show of Word Processors

InDesign 1.0 and 2.0 (Mac and Windows): It handles all OpenType GSUBs flawlessly and it even manages zero-width joiners properly (see K↑h• = K̲h̲ *). It is the only word processor to date that can handle Latin OpenType fonts:

K̲ = K̲	Ľ = Ľ
k̲ = k̲	ľ = ľ
K̲h̲ = K̲h̲ *	Ṁ = Ṁ
k̲h̲ = k̲h̲ *	ṁ = ṁ
K̲H̲ = K̲H̲ *	Ṃ = Ṃ
Ķ = Ķ	ṃ = ṃ
ķ = ķ	M̆ = M̆ *

Word 1997 and 2000: It handles all Unicode characters (on the right of "="), whether PUA or not, but it cannot handle stacking at all (on the left of "="):

A‾• = Ā	E‾• = Ē	H• = Ḥ
a‾• = ā	e‾• = ē	h• = ḥ
A˘• = Ă	E˘• = Ĕ	H• = Ḫ *
a˘• = ă	e˘• = ĕ	h• = ḫ *
A` • = À	E` • = È	Ĥ• = Ĥ
a` • = à	e` • = è	ĥ• = ĥ
A´• = Á	E´• = É	I‾• = Ī
a´• = á	e´• = é	i‾• = ī
A~• = Ã	E~• = Ẽ	I˘• = Ĭ

Word XP: It stacks, if code points are available by Unicode for the respective precomposed characters. PUA glyphs are not stacked. Several bugs, see ȍ etc.:

ḥ = ḥ	L̥ˋ = L̥̀ *	Ōˇ = Ǒ̄ *
Ḥ = Ḥ *	l̥ˋ = l̥̀ *	ōˇ = ō̆ *
ḥ = ḥ *	L̥ˊ = Ḷ́ *	Ōˋ = Ṑ
Ĥ = Ĥ	l̥ˊ = ĺ̥ *	Ò = ò
ĥ = ĥ	L̥ˊ = Ĺ̥ *	Ōˊ = Ó̄
Ī = Ī	l̥ˊ = ĺ̥ *	Ó = ó
ī = ī	L̥˜ = L̥̃ *	Ō˜ = Ȭ *
Ĭ = Ĭ	l̥˜ = ĩ̥ *	ō˜ = ȭ *
ĭ = ĭ	L̥˚ = L̥̊ *	Ő = Ő

TextEdit for Mac: It stacks like Word XP. But it also tries to stack PUA glyphs by combining accents, whether bullet accents (left) or floating accents (right):

Ã = Ã	Ã = Ã
ã = ã	ã = ã
Â = Â	Â = Â
â = â	â = â
Ä = Ä	Ä = Ä
ä = ä	ä = ä
Ā̆ = Ā̆ *	Ā̆ = Ā̆ *
ā̆ = ā̆ *	ā̆ = ā̆ *

ThinkFree for Mac (left): It is Unicode-savvy, but not stacking-savvy at all.
OmniWeb for Mac (right): It is stacking-savvy to the same extent as Word XP:

Ã̰ = Ã	Ã = Ã
ã̰ = ã	ã = ã
Â̰ = Â	Â = Â
â̰ = â	â = â
Ä̰ = Ä	Ä = Ä
ä̰ = ä	ä = ä
Ā̰̆ = Ä̆ *	Ā̆ = Ä̆ *
ā̰̆ = ā̆ *	ā̆ = ā̆ *
Ā̰̀ = Ā̀ *	Ā̀ = Ā̀ *
ā̰̀ = ā̀ *	ā̀ = ā̀ *

WordPerfect 9 (Windows): It is "rubbish-savvy" by defaulting to various fonts.
It seems that it is neither Unicode-savvy nor OpenType-savvy, and it seems
that it can handle only glyphs belonging to code page 1252. I cannot discover
any logical thinking behind the programmer's "font defaulting methods":

m̃ = *	r.` = ׀ *	u´ = ú
N =	Rₐ = ׀ *	Ũ = Ǔ
n =	r₀ = ׀ *	ũ = ũ
Ň = *	R. = ׀ *	Û = Û
ñ = *	r. = ח *	û = û
Ñ = Ñ	R = ׀ *	Ü = Ü
ñ = ñ	r₀ = ׀ *	ü = ü
Ṅ =	R. = ט *	U = ב *
ṅ =	r. = ׳ *	u = ג *

Netscape 7 Macintosh (left): It is Unicode-savvy, but not stacking-savvy at all.
Internet Explorer 6 Windows (right): It is Unicode-savvy and stacking-savvy to the same extent as Word XP, and hence it has the same bugs as Word XP:

g• = ġ	ġ = ġ
G• = Ğ	Ğ = Ğ
g• = ğ	ğ = ğ
G• = Ğ	G• = Ğ
g• = ğ	ğ = ğ
G• = Ĝ	Ĝ = Ĝ

QuarkXPress 5: It is apparently neither Unicode-savvy nor OpenType-savvy. It is interesting to notice that it substitutes combining diacritical marks by spacing marks (both contained in Palladio HOT), e.g. ◌̄ by ¯, if these marks are part of code page 1252, but it is incapable to do so for other diacritical marks, e.g. ◌̆ does not belong to code page 1252, and hence it is substituted by "?":

A¯ = A	e¨ = ë	I¯~ = ? *
a¯ = a	E¯? = ? *	i¯~ = ? *
A? = A	e¯? = ? *	I? = I
a? = a	E^ = ?	i? = i
A` = À	e^ = ?	J? = ? *
a` = à	E´ = ?	j? = j
A´ = Á	e´ = ?	J´ = ? *
a´ = á	E¯~ = ? *	j´ = ? *
A~ = Ã	e¯~ = ? *	J^ = J
a~ = ã	E? = E	j^ = j

VenturaPublisher 8 (Windows): It is apparently neither Unicode-savvy nor OpenType-savvy and behaves like QuarkXPress (see above).

A⁻ = A	e¨ = ë	I⁻~ = ? *
a⁻ = a	E⁻? = ? *	i⁻~ = ? *
A? = A	e⁻? = ? *	I? = I
a? = a	E⁻ˆ = ?	i? = i
A` = À	e⁻ˆ = ?	J? = ? *
a` = à	E⁻´ = ?	j? = j
A´ = Á	e⁻´ = ?	J´ = ? *
a´ = á	E⁻~ = ? *	j´ = ? *
A~ = Ã	e⁻~ = ? *	J^ = J
a~ = ã	E? = E	j^ = j

Palatino Linotype + Word XP (left) versus InDesign (right): The following comparison proves that OpenType features contained in URW Palladio HOT and OpenType features contained in Palatino Linotype are completely ignored by Word XP (left), whereas they are evaluated by InDesign (right):

sub f f i by ffi;	sub f f i by ffi;
sub f f l by ffl;	sub f f l by ffl;
sub f f t by fft;	sub f f t by fft;
sub f f b by ffb;	sub f f b by ffb;
sub f f h by ffh;	sub f f h by ffh;

"Many widely used applications are not savvy"

(OpenType® User Guide for Adobe® Fonts, August 2002, page 7)

8. Storage of Accents in Unicode Text Files

If you enter this test text with **Word 97** (which cannot handle stacking of accents at all)

---

1. E◌́ = É; e◌́ = é; E◌̀ = È; e◌̀ = è

2. E◌̄◌́ = Ḗ; e◌̄◌́ = ḗ; E◌̄◌̀ = Ḕ; e◌̄◌̀ = ḕ

3. G◌̆ = Ğ; g◌̆ = ğ

4. O◌̄◌́ = Ố; o◌̄◌́ = ố; O◌̄◌̀ = Ồ; o◌̄◌̀ = ồ

---

and open it with **Word XP** (which handles stacking of accents imperfectly), you see this:

---

1. É = É; é = é; È = È; è = è

2. Ḗ = Ḗ; ḗ = ḗ; Ḕ = Ḕ; ḕ = ḕ

3. G◌̆ = Ğ; ğ = ğ

4. Ō◌́ = Ố; Ố = ố; Ō◌̀ = Ồ; Ồ = ồ

If this file is saved by Word XP as Unicode TXT file (16-bit, low byte first = "little endian"), it will have the following internal storage structure:

```
HEXEDIT 1.00 von Robert L. Hummel 1992. Hier und da geändert von U. Stiehl
┌──────────┬────────────────────────────────────────────────┬──────────────┐
│ Offset   │                   Hex-Daten                      │  Ascii-Daten │
├──────────┼────────────────────────────────────────────────┼──────────────┤
│ 00000000 │ FF FE 31 00 2E 00 20 00 45 00 01 03 20 00 3D 00 │ ■1 .   E ☺♥ = │
│ 00000010 │ 20 00 C9 00 3B 00 20 00 65 00 01 03 20 00 3D 00 │  ┌ ;   e ☺♥ = │
│ 00000020 │ 20 00 E9 00 3B 00 20 00 45 00 00 03 20 00 3D 00 │  ú ;   E  ♥  =│
│ 00000030 │ 20 00 C8 00 3B 00 20 00 65 00 00 03 20 00 3D 00 │  └ ;   e  ♥  =│
│ 00000040 │ 20 00 E8 00 0D 00 0A 00 32 00 2E 00 20 00 45 00 │  þ ♪ ◙ 2 .   E│
│ 00000050 │ 04 03 01 03 20 00 3D 00 20 00 16 1E 3B 00 20 00 │ ♦♥☺♥  =   ▬▲; │
│ 00000060 │ 65 00 04 03 01 03 20 00 3D 00 20 00 17 1E 3B 00 │ e ♦♥☺♥  =  ↕▲;│
│ 00000070 │ 20 00 45 00 04 03 00 03 20 00 3D 00 20 00 14 1E │  E ♦♥ ♥  =  ¶▲│
│ 00000080 │ 3B 00 20 00 65 00 04 03 00 03 20 00 3D 00 20 00 │ ;   e ♦♥ ♥  = │
│ 00000090 │ 15 1E 0D 00 0A 00 33 00 2E 00 20 00 47 00 06 03 │ §▲♪ ◙ 3 .   G ♠♥│
│ 000000A0 │ 20 00 3D 00 20 00 1E 01 3B 00 20 00 67 00 06 03 │  =   ▲☺;   g ♠♥│
│ 000000B0 │ 20 00 3D 00 20 00 1F 01 0D 00 0A 00 34 00 2E 00 │  =   ▼☺♪ ◙ 4 .│
│ 000000C0 │ 20 00 4F 00 04 03 01 03 20 00 3D 00 20 00 52 1E │  O ♦♥☺♥  =   R▲│
│ 000000D0 │ 3B 00 20 00 6F 00 04 03 01 03 20 00 3D 00 20 00 │ ;   o ♦♥☺♥  = │
│ 000000E0 │ 53 1E 3B 00 20 00 4F 00 04 03 00 03 20 00 3D 00 │ S▲;   O ♦♥ ♥  =│
│ 000000F0 │ 20 00 50 1E 3B 00 20 00 6F 00 04 03 00 03 20 00 │  P▲;   o ♦♥ ♥ │
└──────────┴────────────────────────────────────────────────┴──────────────┘
  Cursor: ←  →  ↑  ↓  Bild↑ Bild↓      F8 = Hex/Ascii    Esc = Sichern/Ende
```

For instance, "E◌́" to the left of "=" is rendered in Word XP as "É" and stored as "45 00 01 03", whereas the precomposed "É" to the right of "=" (also displayed as "É") is stored as "C9 00". So, if you see "É" on screen, you cannot decide, if "É" is stored as "45 00 01 03" or as "C9 00".

---

9. Keyboard Entry of Diacritical Characters

Windows XP users can enter diacritics by these methods (e.g. "ḥ": U+1E25 = 7717 decimal):

1. Num Pad Method (Alt + "0" + <u>decimal</u> Unicode Number)

Diacritics can be entered on the numerical keypad by pressing the Alt key, and while keeping Alt pressed, entering the <u>decimal</u> code on the keypad, e.g. <u>Alt+07717</u> resulting in ḥ. The prefix "0" (zero) may usually be omitted, i.e. Alt+7717.

The num pad method works with Windows XP, but it does not work with Windows 1998, since in Win 98 this method is limited to characters in the decimal range 0-255 (hex 00-FF).

2. Alt-x Method ("U+" + <u>hexadecimal</u> Unicode Number + Alt-x)

Diacritics can be entered by "U+" and the <u>hexadecimal</u> code of the character, followed by Alt-x (i.e. by pressing the keys Alt and "x" simultaneously), e.g. <u>U+1E25 Alt-x</u> resulting in ḥ. It is often possible to omit the prefix "U+", so that here e.g. <u>1E25 Alt-x</u> could be sufficient depending on what letter precedes: If any of the letters "a, b, c, d, e, f" precedes, then the faulty Windows XP parser routine fails and evaluates these letters as hexadecimal digits (e.g. dūtaḥ cannot be entered by <u>d016B Alt-x</u> so that <u>dU+016B Alt-x</u> is required in this case).

The Alt-x method works with Windows XP, but does not work with Windows 98.

> *Note: The Num Pad and Alt-x Method work properly with Wordpad XP. They do not work properly or not at all with Word XP, Notepad XP, InDesign CS, QuarkXpress 6.0, OpenOffice Writer 1.1.*

3. "Eureka" and similar Methods

There exist other entry methods relying on special features of word processing programs, e.g. the "Eureka" method (see http://www.sanskritweb.net/fonts/eureka.doc).

Furthermore it is possible to define keyboard macros (or so-called keyboard shortcuts, see http://www.sanskritweb.net/fonts/pausx.pdf) for the most often used diacritics.

4. "Keyman" Driver

A few third-party developers provide special Windows keyboard drivers, e.g. "Keyman", see http://www.tavultesoft.com.

Furthermore there exist special hardware keyboards with additional freely definable keys.

9.1. Word XP and other Mental Derangements

This refers to MS Windows XP Professional, Version 5.1.2600, Service Pack 1 Build 2600, and to MS Word XP 2002, Version 10.0, Build 4030. Hinweis für Deutsche: Beim deutschen Word XP 2002 ist die untengenannte Tastenkombination nicht Alt-x, sondern Alt-c.

Example: Start Word XP (= Word 2002) on a Windows XP computer, enter "xxx" and mark it up with the Windows XP system font Lucida Sans Unicode. Enter U+F810 Alt-x and you get D', because this font contains D' as Private Use Area code F810 hex = 63504 decimal. However, if you type the seemingly equivalent Alt+63504, you get an arbitrary symbol typeset in the Chinese system font PMingLiU or any other installed CJK font.

U+F810	Alt+63504
D'	盘
Lucida Sans Unicode	PMingLiU

Madness 1 (e.g. U+0904 = 2308 = Devanagari character: ऄ): If you enter U+0904 Alt-x in a line of text marked up with Times New Roman, you do not get the "missing character" symbol of the Times New Roman font, but instead you get an arbitrary Chinese, Japanese or Korean character selected at random by Bill Gates et al. from any of the installed system fonts MingLiU, PMingLiU, SimSun, NSimSun, SimHei, MS Mincho, MS PMincho, MS Gothic, MS PGothic, MS UI Gothic, etc. etc. – It seems that Bill Gates et al. suffer suffer from the delusion that letters unknown to him should be displayed as random Chinese symbols.

Madness 2 (e.g. U+0900 = 2304 = TITUS Cyberbit Basic non-Unicode repha sign): If you enter U+0900 Alt-x in a text marked up with Arial, you do not get the "missing character" symbol of the Arial font, but instead you get the repha sign ˋ of the TITUS Cyberbit Basic font (if installed). – It seems that Bill Gates et al. suffer from the delusion that non-Unicode codes not contained in a font (e.g. Arial) should be displayed by non-Unicode code points contained in any other font (e.g. TITUS).

Madness 3 (e.g. U+F103 = 61699 = URW Palladio HOT PUA character "ȁ"): If you enter Alt+61699 in a line of text marked up with URW Palladio HOT, you do not get the "ȁ" of the URW Palladio HOT font, but instead you get you an arbitrary Chinese, Japanese or Korean character selected at random by Bill Gates et al. – It seems that Bill Gates et al. suffers from the delusion that any PUA codes contained in a font (here URW Palladio HOT) should be displayed as random signs drawn from Chinese fonts.

If the inmates of a madhouse had programmed Word XP, the results could not be worse.