

Alphabetic List of Unicode Latin and IPA Characters

Version 0.41, February 17, 2011 / By Pim Rietbroek



Instruments for Authors

BRILL

1) Purpose of this document

With the implementation of [Unicode](#) on practically every computer sold today we have gained access to a huge character set, which is also robust in that it transfers very well across computer programs and platforms. But finding a way to input a specific character carrying an accent or a diacritic, *h* for instance, is sometimes not so easy. The [character charts](#) found on the [Unicode website](#) are for the most part not arranged alphabetically. And computer manufacturers have at the moment of this writing not supplied a generic input method or keyboard file that gives easy access to all Latin and IPA characters which are to be found in the Unicode Standard. In the absence of such solutions I have compiled an alphabetic list of Latin and IPA characters encoded in the Unicode Standard with their corresponding hexadecimal code point value (hexadecimal – or ‘base-16’ – numbers run from 0 to 9 and then up from A to F).

2) Accents and Diacritics: ‘Precomposed’ versus ‘Decomposed’

In the previous paragraph the word ‘character’ was used loosely to mean both a single letter like *a* and an accented letter like *á*. Without going into the details of the Unicode Standard it is important to note that while many accented letters are encoded as separate so-called ‘precomposed’ characters (like *á*) they are just as validly encoded in an analytic way (‘decomposed’) as ‘base letter plus combining diacritic(s)’, like *a* plus *◌́* (the latter character being ‘combining acute accent’, Unicode hexadecimal 0301; the dotted circle is only displayed here to show that the accent will be combined with the previous character).

One may ask which is the better choice: ‘precomposed’ or ‘decomposed’. The theoretical answer is, ‘neither and both’: as far as the Unicode Standard is concerned they are equivalent. In the practical world of rendering (displaying and/or printing) text, however, some accented letters are rendered more easily using the precomposed glyphs present in fonts, often because there are fonts available which contain these glyphs and keyboard (input) files which give easy access to them. Glyphs like *á* fall into this category. The alphabetical list which forms the core of the current document gives all the Latin characters encoded in Unicode Standard version 5.2, most of them precomposed glyphs. Other glyphs, such as *š*, are not encoded separately in the Unicode Standard and are also absent from most fonts. These must be input as combinations of a base character plus one or more combining diacritics. You must use a so-called ‘[intelligent font](#)’ to get an acceptable rendering of combining diacritics, especially if they are stacked as in *š̂̃̄* (compare this sequence in Times New Roman on Windows XP: *š̂̃̄*).

In short, the guidelines are:

1. If you need a character carrying one or more diacritics, first try to locate it in the list, below.
2. If you cannot locate the glyph in the list, compose the character by adding combining diacritics to the base character.

Following the base character, first start with any diacritic(s) below, working outward from the base character. After that, add any diacritics above, working outward from the base character. To find these essential combining diacritics, [the chart of the Unicode Range Combining Diacritical Marks 0300-036F](#) will prove helpful. Currently, the [Charis SIL](#) fonts are the best choice if you need to use combining diacritics. In the foreseeable future, Brill will release a typeface of its own called 'Brill' (a set of four fonts: roman, italic, bold and bold italic) which covers all of the Latin and IPA ranges and also covers Greek (including classical and biblical Greek and ancient Greek numbers) and Cyrillic (the Slavic part); miscellaneous punctuation marks and symbols used in the humanities complete the Brill character set. The Brill fonts will allow the addition of any number of diacritics to all Latin 'base' characters.

Example of a base letter plus combining diacritics: $\hat{\underset{\cdot}{o}}$ (if this glyph should ever be needed) is encoded as 006F [o], 0323 [$\underset{\cdot}{o}$], 0330 [$\underset{\cdot}{\circ}$], 0304 [\bar{o}], 0302 [\hat{o}].

Technology is moving more and more towards 'decomposed' character input. Successful rendering of 'decomposed' text depends on the capabilities of the text engine of your software and on the complex type technology present in the fonts you use. The more recent your software is, the better chance you have of getting good results.

3) How to input characters by their Unicode hexadecimal number

Having the Unicode value of a character is the first part. The question becomes how to use it. Windows users with access to the MS Office suite from version 2003 can use the 'Alt-X trick': after keying the four-position hex number press Alt-X and the code will be converted to the character or symbol. This works as a toggle: press Alt-X again and the code will reappear. [It does not matter whether you key upper- or lowercase letters.] See the illustrations in [Appendix One](#).

Mac OS X users need to activate the Unicode hex keyboard/input method, which is present in all recent Mac OS X systems. It will work systemwide in all applications. Choose the Unicode hex keyboard, hold down the Option key and key the four-digit hex number, and release the Option key. For Unicode hex keyboard activation, see the illustrations in [Appendix Two](#).

4) Fonts and 'Intelligent' Fonts

Make sure you use a Unicode font. Although there are currently many Unicode fonts in existence, very few exist that offer a wide range of precomposed characters such as *ḥ*, *ṃ*, etc, and are available not only in roman format but also in italic and other styles. One (free) set of fonts which provides both is the [Gentium family](#), available from sil.org; these fonts do not support combining diacritics to any great

extent. SIL also offers the ‘intelligent’ [Charis SIL](#) font family which features good combining diacritics support and which also has a larger character set. [In this document, apart from Charis SIL Unicode, several other fonts have been used, because even Charis SIL does not have all the Latin and characters found in Unicode version 5.2, let alone the current version 6.0.]

5) List of Latin and IPA characters and their Unicode hexadecimal values

a	0061	A	0041	
À	1D00	—	—	(small cap, linguistic use only)
ª	1D43	ˆ	1D2C	(superscript, linguistic use only)
ₐ	2090	—	—	(subscript, linguistic use only)
ⱸ	0250	Ɀ	2C6F	
ᵛ	1D44	—	—	(superscript, linguistic use only)
ɑ	0251	Ɀ	2C6D	
ᵃ	1D45	—	—	(superscript, linguistic use only)
ɒ	0252	D	2C70	
ᵔ	1D9B	—	—	(superscript, linguistic use only)
á	00E1	Á	00C1	
à	00E0	À	00C0	
â	00E2	Â	00C2	
ä	00E4	Ä	00C4	
ã	00E3	Ã	00C3	
ā	0101	Ā	0100	
ǎ	0103	Ǻ	0102	
ǧ	01CE	Ǻ	01CD	
å	00E5	Å	00C5	
à	0201	À	0200	
â	0203	Â	0202	
à	0227	À	0226	
à	1E9A	—	—	
à	1EA3	À	1EA2	
ǻ	2C65	Ǽ	023A	
ą	0105	Ą	0104	
ạ	1EA1	Ạ	1EA0	
ạ	1E01	Ḑ	1E00	
ą	1D8F	—	—	
ą	1D90	—	—	
ǻ	01FB	Ǽ	01FA	
ǻ	1EA5	Ǽ	1EA4	
à	1EA7	À	1EA6	
à	1EA9	À	1EA8	
ã	1EAB	Ã	1EAA	
â	1EAD	Â	1EAC	
ǻ	1EAF	Ǽ	1EAE	
à	1EB1	À	1EB0	

ǎ	1EB3	Ǻ	1EB2	
ǣ	1EB5	ǻ	1EB4	
ǿ	1EB7	Ǽ	1EB6	
ā	01DF	Ā	01DE	
ǻ	01E1	Ǽ	01E0	
æ	00E6	Æ	00C6	
Æ	1D01	—	—	(small cap, linguistic use only)
—	—	Æ	1D2D	(superscript, linguistic use only)
æ	1D02	—	—	
æ	1D46	—	—	(superscript, linguistic use only)
aa	A733	AA	A732	
ao	A735	AO	A734	
aj	A737	AJ	A736	
av	A739	AV	A738	
av	A73B	AV	A73A	
ay	A73D	AY	A73C	
é	01FD	É	01FC	
ē	01E3	Ē	01E2	
b	0062	B	0042	
B	0299	—	—	(small cap, linguistic use only)
b	1D47	B	1D2E	(superscript, linguistic use only)
B	1D03	—	—	(small cap, linguistic use only)
—	—	B	1D2F	(superscript, linguistic use only)
ḃ	1E03	Ḃ	1E02	
Ḅ	0253	Ḃ	0181	
Ḅ	0183	Ḃ	0182	
Ḅ	0180	Ḃ	0243	
Ḅ	1D6C	—	—	
Ḅ	1E05	Ḃ	1E04	
Ḅ	1E07	Ḃ	1E06	
Ḅ	1D80	—	—	
Ḅ	0185	Ḃ	0184	
c	0063	C	0043	
ᶜ	1D9C	—	—	(superscript, linguistic use only)
ɔ	0254	ɔ	0186	(typographically a turned c; lower-mid back rounded vowel, open o)
ɔ	1D10	—	—	(small cap, linguistic use only)
ɔ	1D97	—	—	
ɕ	0297	—	—	(palatal (or alveolar) click; see also ɰ, 01C3)
c	1D04	—	—	(small cap, linguistic use only)
ć	0107	Ć	0106	

ê	0109	Ê	0108		
ě	010D	Ě	010C		
ĉ	010B	Ĉ	010A		
ċ	0188	Ċ	0187		
ȡ	023C	Ȣ	023B		
ḡ	A73F	Ḡ	A73E		
ç	00E7	Ç	00C7		
Ϣ	0255	—	—		
ᵉ	1D9D	—	—	(superscript, linguistic use only)	
ċ̇	1E09	Ĉ̇	1E08		
d	0064	D	0044		
ᵀ	1D05	—	—	(small cap, linguistic use only)	
ᵀ	1D48	ᵀ	1D30	(superscript, linguistic use only)	
δ	1E9F	—	—	(linguistic use, not for Greek)	
g	018D	—	—		
đ	010F	Đ	010E		
ḏ	0111	Ḑ	0110	(compare capital Eth, Ð, 00D0; African D, Ð, 0189)	
ḑ	1D6D	—	—		
ḑ̇	1E0B	Ḑ̇	1E0A		
ḑ̈	0257	Ḑ̈	018A		
ḑ̊	1D91	—	—		
ḑ̋	018C	Ḑ̋	018B		
ḑ̌	A771	—	—	(dum)	
ḑ̍	1E0D	Ḑ̍	1E0C		
ḑ̎	1E0F	Ḑ̎	1E0E		
ḑ̏	1E11	Ḑ̏	1E10		
ḑ̐	1E13	Ḑ̐	1E12		
ḑ̑	0256	Ḑ̑	0189	(compare capital Eth, Ð, 00D0; capital letter D with stroke, Ð, 0110)	
ḑ̒	0221	—	—		
ḑ̓	1D81	—	—		
ḑ̔	0238	—	—		
ḑ̕	02A3	—	—		
ḑ̖	02A5	—	—		
ḑ̗	02A4	—	—		
dz	01F3	Dz	01F2	DZ	01F1
dž	01C6	Dž	01C5	DŽ	01C4
e	0065	E	0045		
ᵉ	1D07	—	—	(small cap, linguistic use only)	
ᵉ	1D49	ᵉ	1D31	(superscript, linguistic use only)	
ₑ	2091	—	—	(subscript, linguistic use only)	

ə	0259	Ə	018F	(Azerbaijani; linguistic use)
ª	1D4A	Ⓐ	1D32	(superscript, linguistic use only)
¸	2094	—	—	(subscript, linguistic use only)
ə	01DD	Ḃ	018E	(Pan-Nigerian alphabet)
Ɑ	2C7B	—	—	(small cap, linguistic use only)
ə̣	1D95	—	—	
ə̤	0258	—	—	(linguistic use)
ə̥	025A	—	—	(linguistic use)
ε	025B	Ε	0190	(linguistic use, not for Greek text)
ε̇	1D4B	—	—	(superscript, linguistic use only)
ε̣	1D93	—	—	
з	1D08	—	—	
³	1D4C	—	—	(superscript, linguistic use only)
Ʒ	1D94	—	—	
з	025C	—	—	
³	1D9F	—	—	(superscript, linguistic use only)
з̇	025D	—	—	
Ɔ	025E	—	—	
ø	029A	—	—	
Ɑ	2C7B	—	—	(small cap, linguistic use only)
é	00E9	É	00C9	
è	00E8	È	00C8	
ê	00EA	Ê	00CA	
ë	00EB	Ë	00CB	
ẽ	1EBD	Ẽ	1EBC	
ē	0113	Ē	0112	
ě	0115	Ě	0114	
ě	011B	Ě	011A	
è	0205	Ẽ	0204	
ê	0207	Ê	0206	
è	0117	Ě	0116	
è	1EBB	Ẽ	1EBA	
e	2C78	—	—	
ø	0247	Ɔ	0246	
ẹ	0119	Ẹ	0118	
ẹ	1EB9	Ẹ	1EB8	
ẹ	0229	Ẹ	0228	
ẹ	1E19	Ẹ	1E18	
ẹ	1E1B	Ẹ	1E1A	
ẹ	1D92	—	—	

é	1E17	É	1E16	
è	1E15	È	1E14	
ċ	1E1D	Ċ	1E1C	
ê	1EBF	Ê	1EBE	
è	1EC1	È	1EC0	
è	1EC3	È	1EC2	
ẽ	1EC5	Ê	1EC4	
ê	1EC7	Ê	1EC6	
f	0066	F	0046	
F	A730	—	—	(small cap, medievalist use only)
f	1DA0	—	—	(superscript, linguistic use only)
f	0192	F	0191	
ƒ	025F	—	—	(turned f, used as Latin small letter dotless j with stroke)
ƒ	1DA1	—	—	(superscript, linguistic use only)
—	—	ƒ	A7FB	
ϕ	0278	—	—	(linguistic use, not for Greek text)
ϕ	1DB2	—	—	(superscript, linguistic use only)
ϖ	2C77	—	—	
ƒ	1E1F	ƒ	1E1E	
ƒ	1D6E	—	—	
f	1D82	—	—	
ƒ	02A9	—	—	
g	0067	G	0047	
g	0261	—	—	
ᵍ	1DA2	—	—	(superscript, linguistic use only)
ǧ	1D79	Ǧ	A77D	
G	0262	—	—	(small cap, linguistic use only)
ᵍ	1D4D	ᵍ	1D33	(superscript, linguistic use only)
g	1D77	—	—	
ɣ	0263	Ƴ	0194	(for African languages and linguistic use, not for Greek text)
ɣ	02E0	—	—	(superscript, linguistic use only)
ɣ	0264	—	—	(Latin small letter rams horn = “Latin small letter baby gamma”)
ǧ	01F5	Ǧ	01F4	
ĝ	011D	Ĝ	011C	
ḡ	1E21	Ḡ	1E20	
ǧ	011F	Ǧ	011E	
ǧ	01E7	Ǧ	01E6	
ǧ	0121	Ǧ	0120	
ǧ	0260	Ǧ	0193	
ǧ	029B	—	—	(small cap, linguistic use only)

ğ	0123	Ğ	0122	
g	01E5	Ḡ	01E4	
g̃	1D83	—	—	
h	0068	H	0048	
ɥ	0265	—	—	
ᵀ	1DA3	—	—	(superscript, linguistic use only)
ɥ	02AE	—	—	
ɥ	02AF	—	—	
Ḥ	029C	—	—	(small cap, linguistic use only)
ḥ	02B0	Ḥ	1D34	(superscript, linguistic use only)
ƒ	2C76	Ƒ	2C75	
ĥ	0125	Ĥ	0124	
ḥ	1E27	Ḥ	1E26	
ħ	021F	Ḥ	021E	
ḥ	1E23	Ḥ	1E22	
ħ	0127	Ḥ	0126	
fi	0266	—	—	
fi	02B1	—	—	(superscript, linguistic use only)
fj	0267	—	—	
h̃	A727	H̃	A726	
ḥ	1E25	Ḥ	1E24	
ḥ	1E29	Ḥ	1E28	
ḥ	2C68	Ḥ	2C67	
ḥ	1E2B	Ḥ	1E2A	
ḥ	1E96	—	—	
hu	0195	Hu	01F6	
i	0069	I	0049	
ᵀ	2071	ᵀ	1D35	(superscript, linguistic use only)
ɪ	026A	(ɪ	0197)	(small cap, linguistic use only)
ᵀ	1DA6	—	—	(superscript small cap, linguistic use only)
ɪ	1D7B	—	—	(small cap, linguistic use only; see also i 0268)
ᵀ	1DA7	—	—	(superscript small cap with stroke, linguistic use only)
ᵀ	1D62	—	—	(subscript, linguistic use only)
ı	1D09	—	—	
ı	1D4E	—	—	
ı	0131	(Turkish etc.)		
İ	0130	(Turkish etc.)		
—	—	İ	A7FE	
ι	0269	ι	0196	(iota: linguistic use, not for Greek text)
ᵀ	1DA5	—	—	(superscript, linguistic use only)

ı	1D7C	—	—	
í	00ED	Í	00CD	
ì	00EC	Ì	00CC	
î	00EE	Î	00CE	
ï	00EF	Ï	00CF	
ĩ	0129	Ĩ	0128	
ī	012B	Ī	012A	
ÿ	012D	ÿ	012C	
ÿ	01D0	Ÿ	01CF	
ÿ	0209	ÿ	0208	
î	020B	Î	020A	
ï	1EC9	Ï	1EC8	
ı	0268	Ɔ	0197	(see also ı 1D7B)
ı̇	1DA4	—	—	(superscript, linguistic use only)
ı̈	012F	İ	012E	
ı̇	1ECB	İ	1ECA	
ı̈	1E2D	İ	1E2C	
ı̊	1D96	—	—	
ı̋	1E2F	İ̋	1E2E	
ij	0133	IJ	0132	
j	006A	J	004A	
Ƶ	1D0A	—	—	(small cap, linguistic use only)
ƶ	02B2	ƶ	1D36	(superscript, linguistic use only)
Ʒ	2C7C	—	—	(subscript, linguistic use only)
ƶ	0237	—	—	
Ʒ	025F	—	—	(turned f, used as Latin small letter dotless j with stroke)
ƶ̇	1DA1	—	—	(superscript, linguistic use only)
ĵ	0135	Ĵ	0134	
ĵ	01F0	—	—	
ƶ	0249	ƶ	0248	
ƶ̈	029D	—	—	
ı̇	1DA8	—	—	(superscript, linguistic use only)
k	006B	K	004B	
Ƶ	1D0B	—	—	(small cap, linguistic use only)
ƶ	1D4F	ƶ	1D37	(superscript, linguistic use only)
ƶ	029E	—	—	
Ƶ	0138	—	—	(Greenlandic <i>kra</i>)
ƶ̇	1E31	ƶ̇	1E30	
ƶ̈	01E9	ƶ̈	01E8	
ƶ̊	0199	ƶ̊	0198	

ķ	A741	Ķ	A740		
ķ	A743	Ķ	A742		
ķ	A745	Ķ	A744		
ķ	1E33	Ķ	1E32		
ķ	1E35	Ķ	1E34		
ķ	0137	Ķ	0136		
ķ	1D84	—	—		
ķ	2C6A	Ķ	2C69		
l	006C	L	004C		
l	029F	—	—	(small cap, linguistic use only)	
l	1DAB	—	—	(superscript small cap, linguistic use only)	
l	02E1	l	1D38	(superscript, linguistic use only)	
l	A747	Ļ	A746		
l	013A	Ļ	0139		
l	013E	Ļ	013D		
l	0140	Ļ	013F		
l	A749	Ļ	A748		
l	0142	Ļ	0141		
l	1D0C	—	—	(small cap, linguistic use only)	
l	019A	Ļ	023D		
l	026B	Ļ	2C62		
l	026C	—	—		
l	2C61	Ļ	2C60		
l	A772	—	—	(lum)	
l	1E37	Ļ	1E36		
l	013C	Ļ	013B		
l	1E3B	Ļ	1E3A		
l	1E3D	Ļ	1E3C		
l	0234	—	—		
l	1D85	—	—		
l	1DAA	—	—	(superscript, linguistic use only)	
l	026D	—	—		
l	1DA9	—	—	(superscript, linguistic use only)	
l	1E39	Ļ	1E38		
l	019B	—	—	(linguistic use, not for Greek text)	
l	1EFB	Ļ	1EFA		
l	02AA	—	—		
l	02AB	—	—		
l	026E	—	—		
l	01C9	Lj	01C8	LJ	01C7

m	006D	M	004D	
—	—	W	A7FD	
ᵐ	1D0D	—	—	(small cap, linguistic use only)
ᵐ	1D50	ᵐ	1D39	(superscript, linguistic use only)
u	026F	U	019C	
ᵘ	1D5A	—	—	(superscript, linguistic use only)
ụ	0270	—	—	
ᵘ	1DAD	—	—	(superscript, linguistic use only)
ᵚ	1D1F	—	—	
—	—	W	A7FF	
ṁ	1E3F	Ṁ	1E3E	
ṁ	1E41	Ṁ	1E40	
ṁ	1D6F	—	—	
ṁ	A773	—	—	(mum)
ṁ	1E43	Ṁ	1E42	
ṁ	0271	Ṁ	2C6E	
ṁ	1DAC	—	—	(superscript, linguistic use only)
ṁ	1D86	—	—	
n	006E	N	004E	
N	0274	—	—	(small cap, linguistic use only)
ⁿ	207F	ⁿ	1D3A	(superscript, linguistic use only)
ⁿ	1DB0	—	—	(superscript small cap, linguistic use only)
ᵐ	1D0E	—	—	(small cap, linguistic use only)
—	—	ᵐ	1D3B	(superscript, linguistic use only)
ñ	0144	Ñ	0143	
ṅ	01F9	Ṇ	01F8	
ñ	00F1	Ñ	00D1	
ň	0148	Ň	0147	
ṅ	1E45	Ṇ	1E44	
ṅ	0149	—	—	(this character is deprecated and its use is strongly discouraged)
ṅ	1D70	—	—	
ṅ	A774	—	—	(num)
ṅ	1E47	Ṇ	1E46	
ṅ	1E49	Ṇ	1E48	
ṅ	0146	Ṇ	0145	
ṅ	1E4B	Ṇ	1E4A	
ṅ	0235	—	—	
ṅ	1D87	—	—	
ṅ	014B	Ṇ	014A	
ṅ	1D51	—	—	(superscript, linguistic use only)

η	0273	—	—	
ⁿ	1DAF	—	—	(superscript, linguistic use only)
ɲ	0272	ɳ	019D	
ɳ	1DAE	—	—	(superscript, linguistic use only)
η	019E	Ŋ	0220	
nj	01CC	Nj	01CB	NJ 01CA
o	006F	O	004F	
o	1D0F	—	—	(small cap, linguistic use only)
◌ ^o	1D52	◌ ^o	1D3C	(superscript, linguistic use only)
◌ _o	2092	—	—	(subscript, linguistic use only)
◌ ^o	1D11	—	—	
œ	1D13	—	—	
◌ ^o	1D12	—	—	
ɔ	0254	Ɔ	0186	(typographically a turned c; lower-mid back rounded vowel, open o)
◌ ^o	1D10	—	—	(small cap, linguistic use only)
◌ ^o	1D53	—	—	(superscript, linguistic use only)
ɔ̥	1D97	—	—	
◌ ^o	1D16	—	—	
◌ ^o	1D54	—	—	(superscript, linguistic use only)
◌ ^o	1D17	—	—	
◌ ^o	1D55	—	—	(superscript, linguistic use only)
ω	0277	—	—	
ó	00F3	Ó	00D3	
ò	00F2	Ò	00D2	
ô	00F4	Ô	00D4	
ö	00F6	Ö	00D6	
õ	00F5	Õ	00D5	
ō	014D	Ō	014C	
ǒ	014F	Ǫ	014E	
ǒ	01D2	Ǫ	01D1	
Ǔ	0151	Ǫ	0150	
ð	020D	Ð	020C	
ô	020F	Ô	020E	
ó	022F	Ó	022E	
ó	1ECF	Ó	1ECE	
ø	01A1	Ø	01A0	
ø	A74D	Ø	A74C	
ø	00F8	Ø	00D8	
ø	01FF	Ø	01FE	
ø	0275	Θ	019F	

◦	1DB1	—	—	(superscript, linguistic use only)
ϑ	A74B	Θ	A74A	
ϑ	2C7A	—	—	
ϱ	01EB	Ϟ	01EA	
̄ϱ	01ED	̄Ϟ	01EC	
ϟ	1ECD	Ϡ	1ECC	
ó	1ED1	Ó	1ED0	
ò	1ED3	Ò	1ED2	
õ	1ED5	Õ	1ED4	
õ	1ED7	Õ	1ED6	
ô	1ED9	Ô	1ED8	
ớ	1EDB	Ớ	1EDA	
ờ	1EDD	Ờ	1EDC	
ở	1EDF	Ở	1EDE	
ỡ	1EE1	Ỡ	1EE0	
ợ	1EE3	Ợ	1EE2	
ố	1E53	Ồ	1E52	
ò	1E51	Ồ	1E50	
ö	022B	Ö	022A	
ö	0231	Ö	0230	
ó	1E4D	Ó	1E4C	
ö	1E4F	Ö	1E4E	
ö	022D	Ö	022C	
œ	0153	Œ	0152	
œ	0276	—	—	(small cap, linguistic use only)
æ	1D14	—	—	
ɒ	01A3	Ɔ	01A2	
∞	A74F	∞	A74E	
8	0223	8	0222	([ou], not the figure 8; linguistic use only)
8	1D15	—	—	(small cap, linguistic use only)
—	—	⁸	1D3D	(superscript, linguistic use only)
p	0070	P	0050	
—	—	Ɔ	A7FC	
p	1D18	—	—	(small cap, linguistic use only)
p	1D56	ᵀ	1D3E	(superscript, linguistic use only)
ḗ	1E55	Ḙ	1E54	
ḙ	1E57	Ḛ	1E56	
ß	01A5	Ɔ	01A4	
Ɔ	1D7D	Ɔ	2C63	
Ɔ	A751	Ɔ	A750	

p	1D71	—	—	
ƀ	1D88	—	—	
ɹ	A753	Ɔ	A752	
ɥ	A755	Ɔ	A754	
q	0071	Q	0051	
q̄	02A0	—	—	
q̇	024B	Q̇	024A	
q̈	A757	Q̈	A756	
ϙ	A759	Q̙	A758	
ϕ	0239	—	—	
r	0072	R	0052	
Ṛ	0280	Ṛ	01A6	(yr: small cap, linguistic, Germanic, and Old Norse use only)
Ṛ	02B3	Ṛ	1D3F	(superscript, linguistic use only)
ɹ	0279	—	—	
Ṛ	02B4	—	—	(superscript, linguistic use only)
Ṛ	1D19	—	—	(small cap, linguistic use only)
Ṛ	1D1A	—	—	(small cap, linguistic use only)
Ṛ	0281	—	—	(small cap, linguistic use only)
Ṛ	02B6	—	—	(superscript small cap, linguistic use only)
Ṛ	1D63	—	—	(subscript, linguistic use only)
ɹ	027E	—	—	
ɹ	A75B	ɹ	A75A	
ɹ	A75D	ɹ	A75C	
Ṛ	0155	Ṛ	0154	
Ṛ	0159	Ṛ	0158	
Ṛ	0211	Ṛ	0210	
Ṛ	0213	Ṛ	0212	
Ṛ	1E59	Ṛ	1E58	
Ṛ	024D	R	024C	
ɹ	1D72	—	—	
ɹ	1D73	—	—	
ɹ	A775	—	—	(rum)
Ṛ	A776	—	—	(rum, small cap)
ɹ	1E5B	Ṛ	1E5A	
Ṛ	1E5F	Ṛ	1E5E	
ɹ	0157	Ṛ	0156	
ɹ	1D89	—	—	
ɹ	027C	—	—	
ɹ	027D	Ṛ	2C64	
ɹ	027F	—	—	

ǀ	027A	—	—	
ǁ	2C79	—	—	
ǂ	027B	—	—	
ǃ	02B5	—	—	(superscript, linguistic use only)
Ǆ	1E5D	Ǆ̄	1E5C	
ǅ	0073	ǅ	0053	
ǆ	A731	—	—	(small cap, linguistic use only)
Ǉ	02E2	—	—	(superscript, linguistic use only)
ǈ	01A8	ǈ	01A7	
ǉ	017F	—	—	
Ǌ	1E9B	—	—	
ǋ	1E9D	—	—	
ǌ	1E9C	—	—	
Ǎ	015B	Ǎ̄	015A	
ǎ	015D	ǎ̄	015C	
Ǐ	0161	Ǐ̄	0160	
ǐ	1E61	ǐ̄	1E60	
Ǒ	1D74	—	—	
ǒ	1E63	ǒ̄	1E62	
Ǔ	015F	Ǔ̄	015E	
ǔ	0219	ǔ̄	0218	
ǖ	023F	ǖ̄	2C7E	
Ǘ	0282	—	—	
Ǚ	1DB3	—	—	(superscript, linguistic use only)
ǚ	1D8A	—	—	
Ǜ	0283	Σ	01A9	(Σ, 01A9, is only used for African languages, not for Greek text)
ǜ	1DB4	—	—	(superscript, linguistic use only)
ǝ	0284	—	—	
Ǟ	0286	—	—	
ǟ	1D8B	—	—	
Ǡ	1D98	—	—	
ǡ	0285	—	—	
Ǣ	01AA	—	—	
ǣ	1E65	ǣ̄	1E64	
Ǥ	1E67	Ǥ̄	1E66	
ǥ	1E69	ǥ̄	1E68	
Ǧ	00DF	Ǧ	1E9E	
ǧ	0074	ǧ	0054	
Ǩ	1D1B	—	—	(small cap, linguistic use only)
ǩ	1D57	ǩ̄	1D40	(superscript, linguistic use only)

ɿ	0287	—	—	(dental click; see also , 01C0)
İ	1E97	—	—	
ƚ	0165	ƚ̇	0164	
ı	1E6B	İ	1E6A	
ƒ	01AD	T	01AC	
ƒ	0167	F	0166	
ƒ	1D75	—	—	
ƒ	2C66	ƒ̇	023E	
ƒ	A777	—	—	(tum)
ƒ	0163	ƒ̇	0162	
ƒ	1E6D	ƒ̇	1E6C	
ƒ	1E6F	ƒ̇	1E6E	
ƒ	021B	ƒ̇	021A	
ƒ	1E71	ƒ̇	1E70	
ƒ	0236	—	—	
ƒ	01AB	—	—	
ƒ	1DB5	—	—	(superscript, linguistic use only)
ƒ	0288	ƒ̇	01AE	
ƒ	02A8	—	—	
ƒ	1D7A	—	—	
ƒ	02A6	—	—	
ƒ	02A7	—	—	
ƒ	A729	ƒ̇	A728	(Mayanist use)
u	0075	U	0055	
U	1D1C	—	—	(small cap, linguistic use only)
U	1DB8	—	—	(superscript small cap, linguistic use only)
U	1D7E	—	—	(small cap, linguistic use only)
U	1D58	U	1D41	(superscript, linguistic use only)
U	1D64	—	—	(subscript, linguistic use only)
U	1D1D	—	—	
U	1D59	—	—	(superscript, linguistic use only)
U	028A	U	01B1	
U	1DB7	—	—	(superscript, linguistic use only)
U	1D7F	—	—	
ú	00FA	Ú	00DA	
ù	00F9	Û	00D9	
û	00FB	Û	00DB	
ü	00FC	Û	00DC	
U	1D1E	—	—	
ü	0169	Û	0168	

ū	016B	Ū	016A	
ů	016D	Ů	016C	
ǔ	01D4	Ǔ	01D3	
ũ	016F	Ũ	016E	
ú	0171	Ú	0170	
ù	0215	Ù	0214	
û	0217	Û	0216	
ú	1EE7	Ũ	1EE6	
ư	01B0	Ư	01AF	
ⱸ	0289	ⱷ	0244	
ᵘ	1DB6	—	—	(superscript, linguistic use only)
ɥ	0173	Ƴ	0172	
ɷ	1E73	Ƴ	1E72	
ɸ	1E75	Ƴ	1E74	
ɹ	1E77	Ƴ	1E76	
ɺ	1EE5	Ƴ	1EE4	
ɻ	1D99	—	—	
ü	01D6	Û	01D5	
ú	01D8	Ú	01D7	
ǔ	01DA	Ǔ	01D9	
ù	01DC	Ù	01DB	
ú	1E79	Ú	1E78	
ü	1E7B	Û	1E7A	
ư	1EE9	Ư	1EE8	
ừ	1EEB	Ừ	1EEA	
ử	1EED	Ử	1EEC	
ữ	1EEF	Ữ	1EEE	
ự	1EF1	Ự	1EF0	
ⱸ	1D6B	—	—	
v	0076	V	0056	
v	1D20	—	—	(small cap, linguistic use only)
ᵛ	1D5B	ᵛ	2C7D	(superscript, linguistic use only)
ᵱ	1D65	—	—	(subscript, linguistic use only)
Λ	028C	Λ	0245	
^	1DBA	—	—	(superscript, linguistic use only)
ⱸ	028B	ⱷ	01B2	
ᵐ	1DB9	—	—	(superscript, linguistic use only)
ⱸ	1EFD	ⱷ	1EFC	
ᵱ	2C71	—	—	
ᵱ	1E7D	ᵱ	1E7C	

ʏ	A75F	Ÿ	A75E	
Ƶ	1E7F	ƶ	1E7E	
Ʒ	2C74	—	—	
ƹ	1D8C	—	—	
Ʒ	A761	ƹ	A760	(vy)
w	0077	W	0057	
w	1D21	—	—	(small cap, linguistic use only)
w	02B7	w	1D42	(superscript, linguistic use only)
ʌ	028D	—	—	
ŵ	1E83	Ŵ	1E82	
Ẁ	1E81	Ẃ	1E80	
ŵ	0175	Ŵ	0174	
Ẅ	1E85	Ẃ	1E84	
Ẁ	1E98	—	—	
ẁ	1E87	Ẃ	1E86	
w	2C73	W	2C72	
Ẅ	1E89	Ẃ	1E88	
w̄	02AC	—	—	
x	0078	X	0058	
x	02E3	—	—	(superscript, linguistic use only)
x	2093	—	—	(subscript, linguistic use only)
ẍ	1E8D	Ẍ	1E8C	
ẋ	1E8B	Ẍ	1E8A	
ɣ	1D8D	—	—	
y	0079	Y	0059	
Ʒ	028F	—	—	(small cap, linguistic use only)
Ƶ	02B8	—	—	(superscript, linguistic use only)
ʌ	028E	—	—	
ý	00FD	Ý	00DD	
ÿ	1EF3	Ỳ	1EF2	
ÿ	0177	Ỳ	0176	
ÿ	00FF	Ỳ	0178	
ÿ	1EF9	Ỳ	1EF8	
ȳ	0233	Ȳ	0232	
ÿ	1E99	—	—	
ÿ	1E8F	Ỳ	1E8E	
ÿ	1EF7	Ỳ	1EF6	
Ʒ	01B4	ƹ	01B3	
Ʒ	024F	ƹ	024E	
Ʒ	1EF5	ƹ	1EF4	

Ʒ	1EFF	Ʒ	1EFE	
ij	0133	IJ	0132	
z	007A	Z	005A	
z	1D22	—	—	(small cap, linguistic use only)
z	1DBB	—	—	(superscript, linguistic use only)
ŷ	A763	ŷ	A762	(Visigothic z)
ž	017A	Ž	0179	
ž	1E91	Ž	1E90	
ž	017E	Ž	017D	
ž	017C	Ž	017B	
z	01B6	Z	01B5	
z	1D76	—	—	
ẓ	1E93	Ẓ	1E92	
z̤	1E95	Z̤	1E94	
z	0225	Z	0224	
z	0240	Z	2C7F	
z	0291	—	—	
ẓ	1DBD	—	—	(superscript, linguistic use only)
ẓ	0290	—	—	
ẓ	1DBC	—	—	(superscript, linguistic use only)
ẓ	1D8E	—	—	
z	2C6C	Z	2C6B	
ij	0133	IJ	0132	
ƀ	00DF	ƀ	1E9E	
Ɔ	00FE	Ɔ	00DE	
Ɔ̣	A765	Ɔ̣	A764	
Ɔ̣	A767	Ɔ̣	A766	
ð	00F0	Ð	00D0	(eth; compare capital letter D with stroke, Ð, 0110; African D, Ð, 0189)
ð	1D06	—	—	(small cap, linguistic use only)
ð	1D9E	—	—	(superscript, linguistic use only)
Ʒ	021D	Ʒ	021C	(yogh)
Ʒ	01BF	Ʒ	01F7	(wynn)
Ʒ	A769	Ʒ	A768	(Old Norse vend)
Ʒ	0292	Ʒ	01B7	(ezh)
Ʒ	1D23	—	—	(small cap, linguistic use only)
Ʒ	1DBE	—	—	(superscript, linguistic use only)
Ʒ	01B9	Ʒ	01B8	
ž	01EF	Ž	01EE	
ž	0293	—	—	
ž	01BA	—	—	

̃	1D9A	—	—	
3	A76B	3	A76A	(et)
ƒ	A76D	ƒ	A76C	(is)
9	A76F	9	A76E	(con)
⁹	A770	—	—	(modifier letter us)
Ŷ	A778	—	—	(um)
ʔ	1D24	—	—	
ʔ	0294	—	—	(glottal stop; a caseless letter; casing variants: ʔ, 0242, and ʔ, 0241)
ʔ	0242	ʔ	0241	(glottal stop; casing letters, used in Canadian aboriginal orthographies)
ɿ	0295	—	—	
ᶞ	02E4	—	—	(superscript version of 0295, linguistic use only)
ɬ	0296	—	—	(lateral click; see also ɬ, 01C1)
ʔ	02A1	—	—	
ɤ	02A2	—	—	
ʂ	01BE	—	—	(archaic phonetic for [ts] affricate; inverted glottal stop with stroke)
ʔ	02C0	—	—	(superscript, linguistic use only)
ᶞ	02C1	—	—	(superscript, linguistic use only)
›	02BE	—	—	
ᶜ	02BF	—	—	
ⲓ	A723	ⲓ	A722	(Egyptological alef)
ᶜ	A725	ᶜ	A724	(Egyptological ain)
ʼ	A78C	ʼ	A78B	(saltillo)
‘	02BB	—	—	(Hawai‘ian ‘okina)
•	1D25	—	—	
ᶞ	1D5C	—	—	(superscript, linguistic use only)
𐀀	02AD	—	—	
	01C0	—	—	(dental click; see also ɿ, 0287)
	01C1	—	—	(lateral click; see also ɬ, 0296)
‡	01C2	—	—	(alveolar click; see also ɥ, 0297)
!	01C3	—	—	(retroflex click; see also ɥ, 0297)
⊙	0298	—	—	(bilabial click)
2	01BB	—	—	(two with stroke: archaic phonetic for [dz] affricate)
ɛ	A72B	ɛ	A72A	(tresillo: Mayanist use)
4	A72D	4	A72C	(cuatrillo: Mayanist use)
4	A72F	4	A72E	(cuatrillo with comma: Mayanist use)
5	01BD	5	01BC	(tone five)

Appendix I: Input Unicode hex in MS Office (Windows)

Version 0.1, November 23, 2009 / By Pim Rietbroek



Instruments for Authors

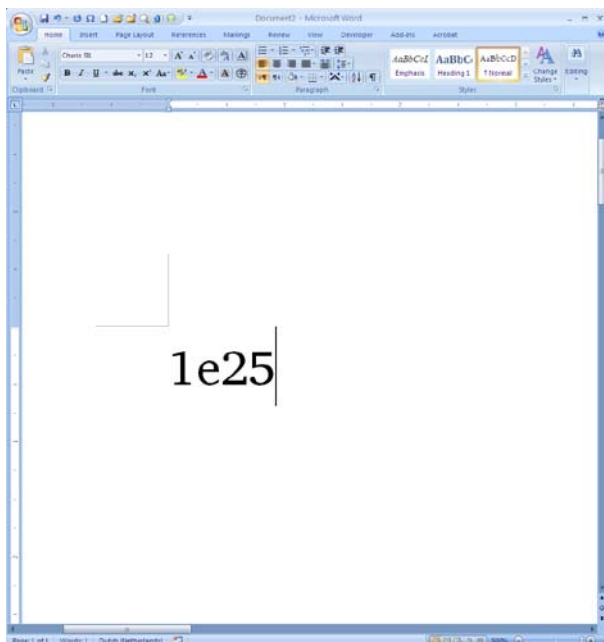
BRILL

How to input characters by their Unicode hexadecimal number

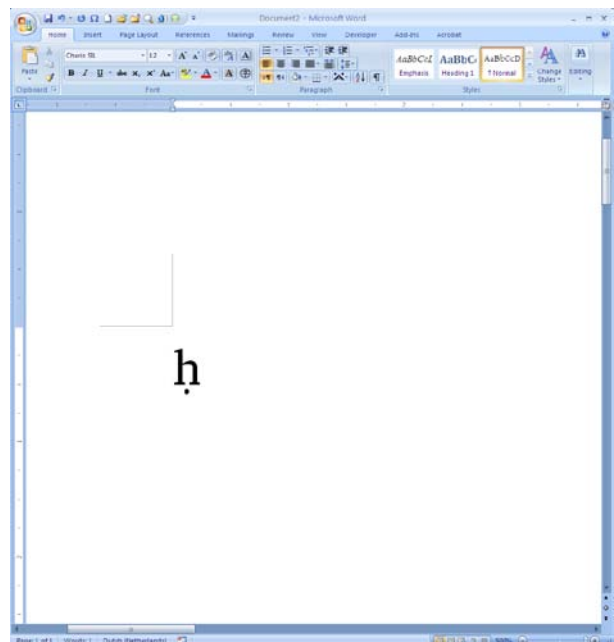
Having the Unicode value of a character is the first part. The question is how to use it.

Windows users with access to the MS Office suite from version 2003 can use the 'Alt-X trick'.

1) Key the four-position hex number (letters may be upper- or lowercase):



2) Press Alt-X and the code will be converted to the character or symbol:



This works as a toggle: press Alt-X again and the code will reappear.

Appendix II: Activate the Unicode Hex Keyboard (Mac OS X)

Version 0.1, November 23, 2009 / By Pim Rietbroek



Instruments for Authors

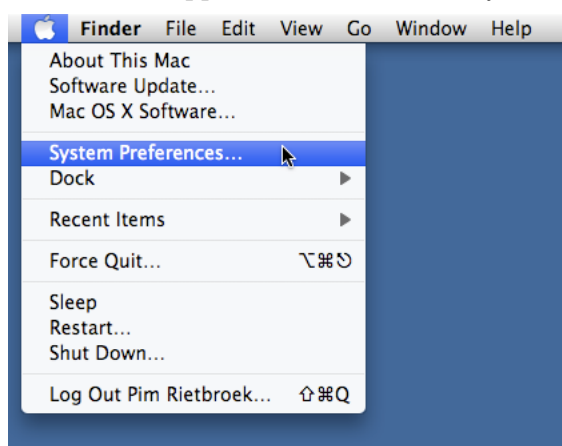
BRILL

How to input characters by their Unicode hexadecimal number

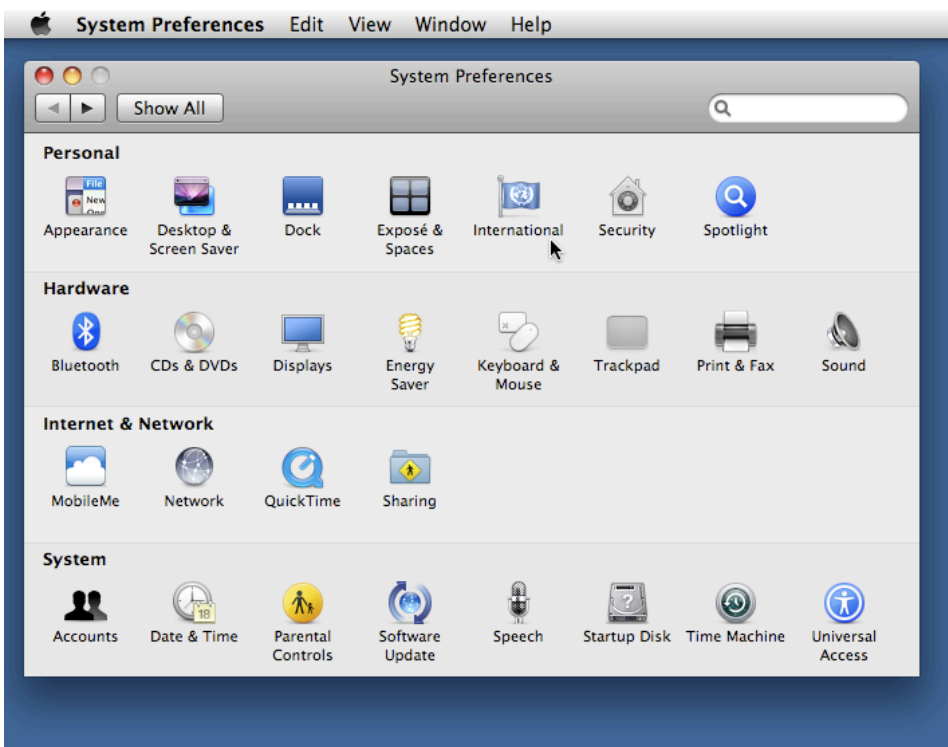
Having the Unicode value of a character is the first part. The question is how to use it.

You need to *activate* the Unicode hex keyboard/input method, which is present on all recent Mac OS X versions (some Snow Leopard [Mac OS X v.6.x] captions are slightly different).

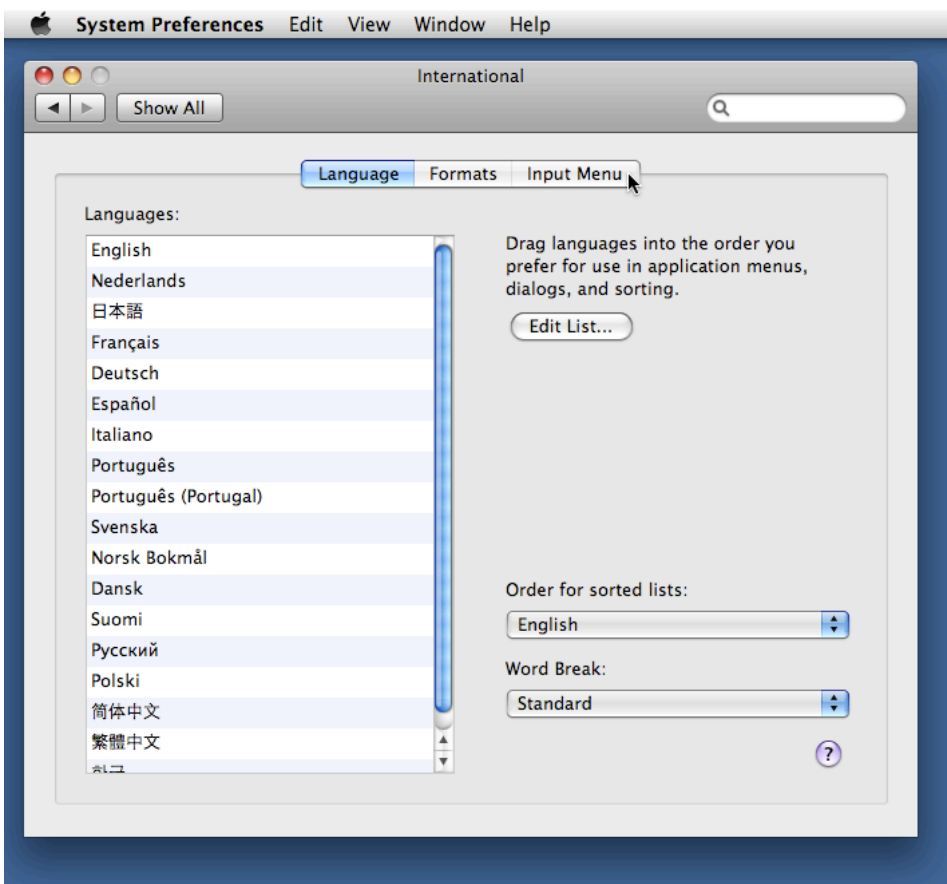
1) Go to the Apple menu and choose System Preferences:

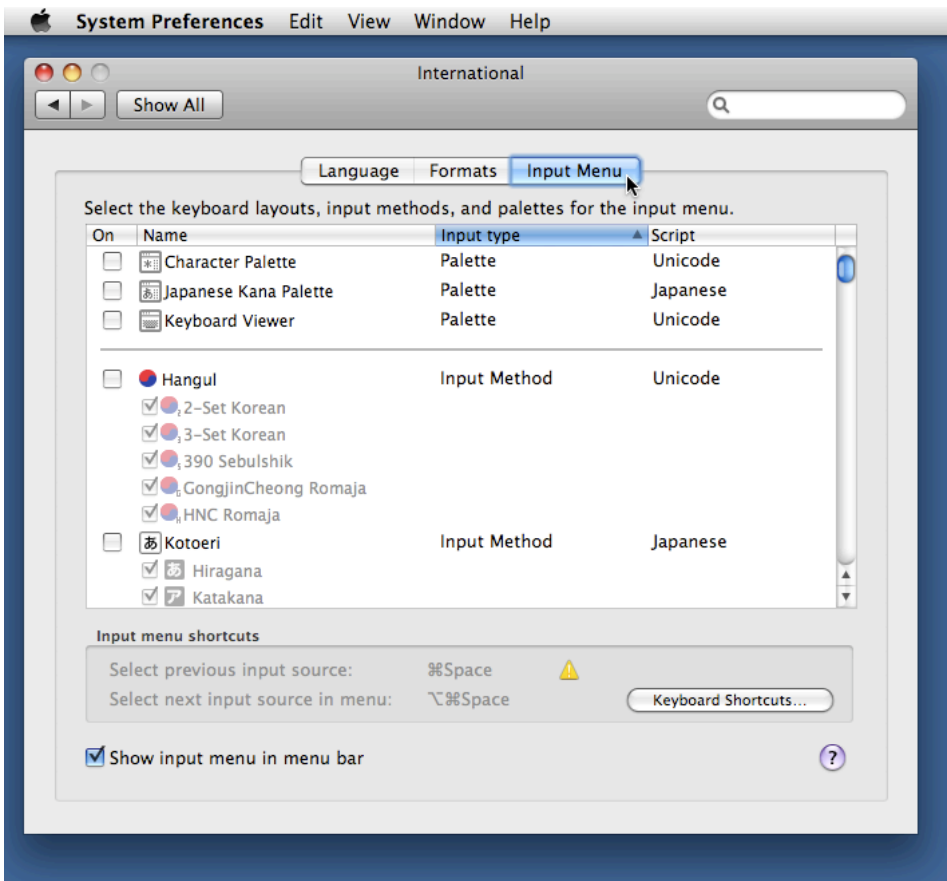


2) In System Preferences, under Personal, click on International (Snow Leopard: Language & Text):

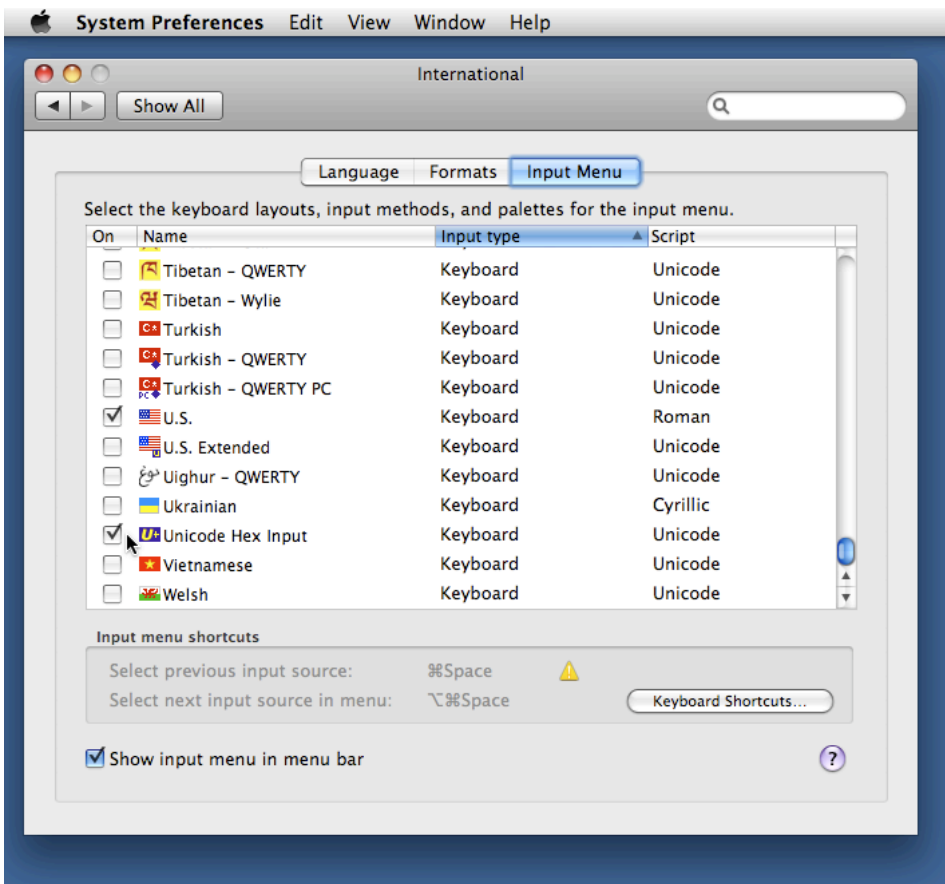


3) Click on the Input Menu button (Input Sources in Snow Leopard):

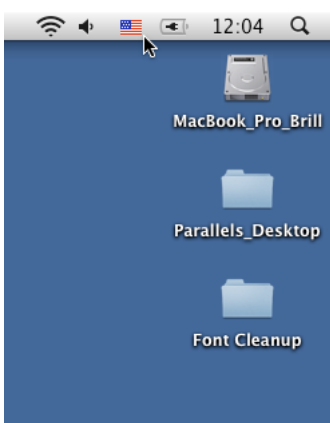




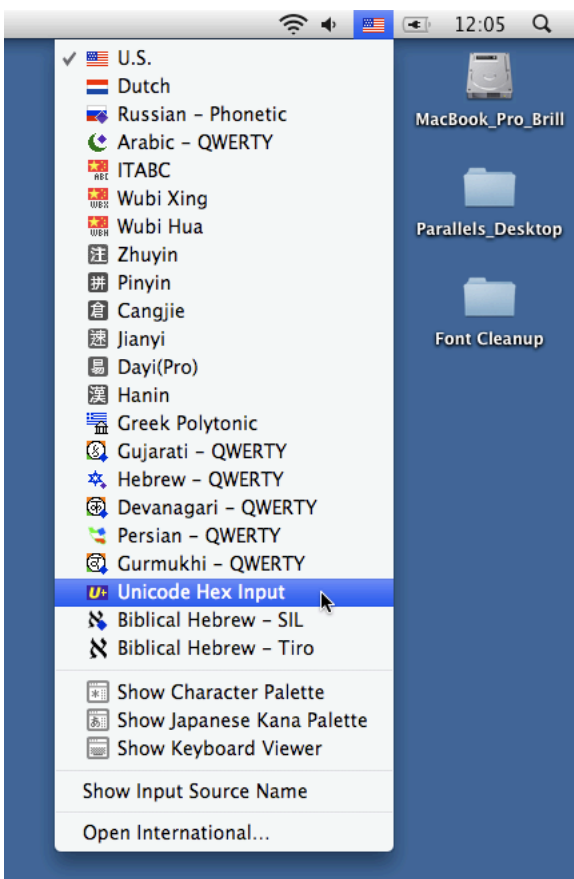
4) Scroll down until you see “Unicode Hex Input” and checkmark “on”. Make sure the option “Show input menu in menu bar” at bottom left (bottom right in Snow Leopard) is checkmarked.



5) Close System Preferences. If it was not visible already a keyboard menu appears in the menu bar near the right-hand side.



6) In order to choose the Unicode hex keyboard, click on the flag symbol and scroll down:



From now on, this keyboard is available system-wide, in all applications that handle text. You use this keyboard to input the four-position hex codes: press down the Option (/Alt) key, type the four-position code, and release the Option key; the character or symbol will then appear at the insertion point. [You can type unaccented text when you leave the Unicode hex keyboard active *without* depressing the Option key.]

It does not matter whether you key upper- or lowercase letters in a hexadecimal number.