Typing combining marks in Junicode 2

Junicode has better support for combining marks (including diacritics and marks of abbreviation) than any other medieval font. You can place a combining mark on just about any character where it makes any sense at all to place it. Just type the base character followed by the mark.

But typing a mark may be a challenge for many users. (I'm not talking about the common letter+diacritic combinations like **á**, for which your system no doubt provides special keyboard sequences, but about exotic combinations like **þ̉** or **m̋**. Stuff that comes up with distressing regularity when you're transcribing medieval manuscripts, whose scribes weren't constrained by the limitations of mechanical typesetting.) Most marks have no assigned keys on a keyboard. Instead you need to know their *encodings*: the four-digit hexadecimal (base-16) number assigned to them in the Unicode standard (for pointers on entering these, see "Getting Started with Junicode" in the *Feature Reference*.

If what I've just written looks like gibberish, don't worry: this document will lay out all the encodings for combining defined by Unicode and likely to be of interest to medievalists. For those not defined by Unicode, the Medieval Unicode Font Initiative (MUFI) has assigned code points, but because these code points are not recognized by Unicode, most applications won't recognize them as combining marks or position them correctly over their base characters: they may come out looking like **m̊**, or worse yet **Å**.

To get around this problem, Junicode provides two alternative methods of entering non-Unicode marks. The first is a collection of **entities**, or mnemonic codes: these are enabled by Stylistic Set 10 (**ss10**), which should be applied to the whole of any document that uses them. These entities avoid the use of MUFI's custom code points and instead associate them with Unicode code points so that applications will recognize them as combining marks.

Junicode's entities will look familiar to anyone who has worked with HTML (the language of web pages): they are preceded by an ampersand (**&**) and followed by a semicolon (**;**), and they consist of the underscore character followed by one or more (mostly mnemonic) letters. For example, to enter the sequence **g̊**, type **g&_oslash;**.

There's not much more to these entities, except for one very important caveat, namely that they affect the appearance of your document but not the underlying text. When you type the sequence **p&_ru;**, it *looks* like **p** with combining rum-sign (**p̛**), but underneath, the text is

still **p&_ru;**. So if you send your text file to a friend who doesn't have Junicode 2 installed, that friend will see the entities, not the marks; if you copy your text onto the clipboard and paste it into an application that either can't access Junicode 2 or can't apply OpenType features, you'll also see the entities. If you are sending your text to a publisher, turn off ss10 to make the entities visible and supply a key explaining what they are. Send along a copy of Junicode, while you're about it, in case your publisher needs a fallback font.

The second method for entering combining marks, for those using applications that support it (Word unfortunately does not), is to apply the OpenType feature **cv84** with the appropriate index to the base character and combining macron (U+0304). This method will be preferable to users who need to preserve searchability—publishers of online texts, for example. For details concerning **cv84**, see the *Feature Reference*, section I; the relevant indices are listed below.

One last very important note: in most text-processing applications, the features needed to make combining marks work are on by default, but *in Microsoft Word they are not*. You've got to turn them on. To do so, open the "Font" dialog, click over to the "Advanced" tab, and enable **Kerning**, **Standard Ligatures**, and **Contextual Alternates**. For entities, you must also enable **Stylistic Set 10** (ss10).

In the following list, non-Unicode marks are highlighted in green.

Encoding	Entity	cv84	Example	Encoding	Entity	cv84	Example
035B	none	-	◌ͤ	F03D	&_thorn;	27	◌ᵬ
1DD3	none	-	◌ᷓ	1DE3	none	-	◌ᷢ
0305	none	-	◌̅	1DE5	none	-	◌ᷥ
0363	none	-	◌ᵃ	1DD8	none	-	◌ᷘ
1DE8	none	-	◌ᵇ	1DD5	none	-	◌ᷝ
0368	none	-	◌ᶜ	1DD6	none	-	◌ᷞ
0369	none	-	◌ᵈ	F135	&_eogo;	14	◌ᵉ
0364	none	-	◌ᵉ	F136	&_emac;	13	◌ᵉ̄

Encoding	Entity	cv84	Example	Encoding	Entity	cv84	Example
1DEB	none	-		F02F	&_idotl;	15	
1DDA	none	-		F031	&_jdotl;	17	
036A	none	-		F13E	&_oogo;	21	
0365	none	-		F032	&_oslash;	22	
F030	&_j;	16		F13F	&_omac;	20	
1DD3	none	-		1DD2	none	-	
1DDD	none	-		1DD1	none	-	
036B	none	-		F013	&_bsc;	11	
1DE0	none	-		F016	&_dsc;	12	
0366	none	-		1DDB	none	-	
1DEE	none	-		F01C	&_ksc;	18	
F033	&_q;	23		1DDD	none	-	
036C	none	-		036B	none	-	
1DE4	none	-		1DE0	noner	-	
036D	none	-		036C	none	-	
0367	none	-		F02A	&_tsc;	25	
036E	none	-		F036	&_an;	7	
1DF1	none	-		F038	&_ar;	9	
036F	none	-		F03A	&_ansc;	8	
F02B	&_y;	26		F130	&_arsc;	10	
1DE6	none	-		F03E	&_orr;	28	

Encoding	Entity	cv84	Example	Encoding	Entity	cv84	Example
1DD4	none	-	◌ᷔ	F03F	&_oru;	29	◌᷒
1DD9	none	-	◌ᷙ	1ACE	none	-	◌᷎