Appendix B: Two-Part Characters

Volume 1 lists composite characters that are formed from a base character and succeeding non-spacing marks that are bound to that base character. There is another related class of characters which also have two or more graphically distinct elements called "multi-part characters." The Devanagari two-part vowels are examples of these characters (see Volume 1, Section 3.1, General Scripts, Devanagari). As with composite characters, there are alternate Unicode character sequences that can be used to express equivalent spellings. These sequences are similar to composite letters such as "â," which can either be spelled with "a" followed by non-spacing "^", or spelled with a single Unicode character "â".

In rendering, the multi-part characters may also behave as if they were separate characters. This occurs, for example, when Indic two-part vowels can have one portion reordered in front of a conjunct consonant. (See Tamil Shaping Behavior, Two-Part vowels in Appendix A.) It may also occur if the text has a different tracking or letter spacing. The following is a list of multi-part characters included in version 1.0 of the Unicode standard.

Unicode value	Part1	Part2	Part3	Name
0132	0049	004A		LATIN CAPITAL LETTER IJ
0133	0069	006A		LATIN SMALL LETTER IJ
01C4	0044	017D		LATIN CAPITAL LETTER DZ HACEK
01C5	0044	017E		LATIN LETTER CAPITAL D SMALL Z HACEK
01C6	0064	017E		LATIN SMALL LETTER DZ HACEK
01C7	004C	004A		LATIN CAPITAL LETTER LJ
01C8	004C	006A		LATIN LETTER CAPITAL L SMALL J
01C9	006C	006A		LATIN SMALL LETTER LJ
01CA	004E	004A		LATIN CAPITAL LETTER NJ
01CB	004E	006A		LATIN LETTER CAPITAL N SMALL J
01CC	006E	006A		LATIN SMALL LETTER NJ
09CB	09C7	09BE		BENGALI VOWEL SIGN O
09CC	09C7	09D7		BENGALI VOWEL SIGN AU
0B48	0B47	0B56*		ORIYA VOWEL SIGN AI
0B4B	0B47	0B3E		ORIYA VOWEL SIGN O
0B4C	0B47	0B57		ORIYA VOWEL SIGN AU
0B94	0BD7	0B92		TAMIL LETTER AU
0BCA	0BC6	0BBE		TAMIL VOWEL SIGN O
0BCB	0BC7	0BBE		TAMIL VOWEL SIGN OO
0BCC	0BC6	0BD7		TAMIL VOWEL SIGN AU
0C48	0C46	0C56		TELUGU VOWEL SIGN AI
0CC0	0CBF	0CD5		KANNADA VOWEL SIGN II
0CC7	0CC6	0CD5		KANNADA VOWEL SIGN EE
0CC8	0CC6	0CD6		KANNADA VOWEL SIGN AI
0CCA	0CC6	0CC2		KANNADA VOWEL SIGN O
0CCB	0CC6	0CC2	0CD5	KANNADA VOWEL SIGN OO
0D4A	0D46	0D3E		MALAYALAM VOWEL SIGN O
0D4B	0D47	0D3E		MALAYALAM VOWEL SIGN OO
0D4C	0D46	0D57		MALAYALAM VOWEL SIGN AU
0E33	0E4D	0E32		THAI VOWEL SIGN SARA AM
0EB3	0ECD	0EB2		LAO VOWEL SIGN AM

* By analogy with the other scripts, this would be the code position for the second part
(U+ 0B56 ORIYA AI LENGTH MARK), but that character was mistakenly omitted from Volume 1.