RESPONSE TO IP LETTER from August, the 19th on Greek-Latin cross-script variants

Latin GP has reviewed the table sent by IP on 9th August 2019. Here are Latin GP’s initial conclusions and remarks:

1. Of the cross-script variants that the Greek GP is proposing, the following have also been identified by the Latin GP:

|  |  |  |
| --- | --- | --- |
| **Greek Letter** | **Latin Letter** | **Latin GP comment** |
| U+03B3 γ | ɣ U+0263 |  Although it is not explicitly defined as variant, it will be due to Transitivity principle (Variants due to underlining) |
| Greek small letter GAMMA | Latin small letter GAMMA |
| U+03B5 ε | ɛ U+025B |   |
| Greek small letter EPSILON | Latin small letter OPEN E |
| U+03B9 ι | ɩ U+0269 |   |
| Greek small letter IOTA | Latin small letter IOTA |
| U+03BD ν | v U+0076 |   |
| Greek small letter NU | Latin small letter V |
| U+03BF ο | o U+006F |   |
| Greek small letter OMICRON | Latin small letter O WITH DOT BELOW |
| U+03BF ο | ọ U+1ECD | Although it is not explicitly defined as variant, it will be due to Transitivity principle (Variants due to underlining) |
| Greek small letter OMICRON | Latin small letter O with dot below |
| U+03C5 υ | ʋ U+028B |   |
| Greek small letter UPSILON | Latin small letter V WITH HOOK |
| U+03AF ί | í U+00ED |   |
| Greek small letter IOTA WITH TONOS | Latin small letter I WITH ACUTE |
| U+03CC ό | ó U+00F3 |   |
| Greek small letter OMICRON WITH TONOS | Latin small letter O WITH ACUTE |
| U+03CA ϊ | ï U+00EF |   |
| Greek small letter IOTA WITH DIALYTIKA | Latin small letter I WITH DIAERESIS |
| U+03CB ϋ | ü U+00FC |   |
| Greek small letter UPSILON WITH DIALYTIKA | Latin small letter U WITH DIAERESIS |

(2) In addition, these cross-script variants (not listed above) have been identified by the Latin GP:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Source Unicode Name** | **Source Code Point** | **Source Glyph** | **Variant Relationship** | **Target Glyph** | **Target Code Point** | **Target Unicode Name** | **Disposition** | **Rationale** |
|  | LATIN SMALL LETTER DOTLESS I | 0131 | ı | ↔ | ι | 03B9 | GREEK SMALL LETTER IOTA | Blocked | Homoglyph |
|  | LATIN SMALL LETTER A | 0061 | a | ↔ | α | 03B1 | GREEK SMALL LETTER ALPHA | Blocked | Glyphs nearly identical due to font design. See **[G1]** below. |
|  | LATIN SMALL LETTER P | 0070 | p | ↔ | ρ | 03C1 | GREEK SMALL LETTER RHO | Blocked | Glyphs nearly identical due to font design. See **[G2]** below. |
|  | LATIN SMALL LETTER U | 0075 | u | ↔ | υ | 03C5 | GREEK SMALL LETTER UPSILON | Blocked | Glyphs nearly identical due to font design. See **[G3]** below. |
|  | LATIN SMALL LETTER Y | 0079 | y | ↔ | γ | 03B3 | GREEK SMALL LETTER GAMMA | Blocked | Glyphs nearly identical due to font design |
|  | LATIN SMALL LETTER SHARP S | 00DF | ß | ↔ | β | 03B2 | GREEK SMALL LETTER BETA | Blocked | Glyphs nearly identical due to font design. See **[G4]** below. |
|  | LATIN SMALL LETTER A WITH ACUTE | 00E1 | á | ↔ | ά | 03AC | GREEK SMALL LETTER ALPHA WITH TONOS | Blocked | Glyphs nearly identical due to font design |
|  | LATIN SMALL LETTER U WITH ACUTE | 00FA | ú | ↔ | ύ | 03CD | GREEK SMALL LETTER UPSILON WITH TONOS | Blocked | Glyphs nearly identical due to font design. See **[G3]** below. |
|  | LATIN SMALL LETTER O WITH HORN | 01A1 | ơ | ↔ | σ | 03C3 | GREEK SMALL LETTER SIGMA | Blocked | Glyphs nearly identical due to font design. See **[G5]** below. |

(3) The following cross-script variants, which were identified by the Greek GP, were not previously identified by the Latin GP and will need to be evaluated by Latin GP:

|  |  |
| --- | --- |
| **Greek Letter** | **Latin Letter** |
| U+03B7 ηGreek small letter ETA | n U+006ELatin small letter N |
| U+03B7 ηGreek small letter ETA | ņ U+0146Latin small letter N WITH CEDILLA |
| U+03B7 ηGreek small letter ETA | ŋ U+014BLatin small letter ENG |
| U+03B7 ηGreek small letter ETA | ṇ U+1E47Latin small letter N WITH DOT BELOW |
| U+03AE ήGreek small letter ETA WITH TONOS | ń U+0144Latin small letter N WITH ACUTE |
| U+03AF ίGreek small letter IOTA WITH TONOS | i U+0069Latin small letter I |
| U+03AF ίGreek small letter IOTA WITH TONOS | ì U+00ECLatin small letter I WITH GRAVE |
| U+03AF ίGreek small letter IOTA WITH TONOS | ò U+00F2Latin small letter O WITH GRAVE |
| U+03CC όGreek small letter OMICRON WITH TONOS | ȯ U+022FLatin small letter O WITH DOT ABOVE |

(4) These two cases, identified by the Greek GP as variants, involve code points which are not in the Latin repertoire. Accordingly, they can be dropped. (See also the PS below.)

|  |  |
| --- | --- |
| **Greek Letter** | **Latin Letter** |
| U+03B1 α | ɑ U+0251 |
| Greek small letter ALPHA | Latin small letter ALPHA |
| U+03BD ν | ṿ U+1E7F |
| Greek small letter NU | Latin small letter V WITH DOT BELOW |

The IP notes that

Some of the proposed cross-script variants would, if accepted, impose in-script variants for the Latin script. We would like to appraise the Latin GP of these and would like to get the GP's reaction to these proposals.

In principle, each LGR defines its own set of variants, appropriate to its user community, and integration results in the common superset. However, when this process would result in imposing in-script variants between two characters that are both in another script, the bar for accepting such variants for integration has to be necessarily rather higher.

Regarding possible in-script variants as a result of the Greek/Latin cross-script variants:

For most of the cases (those involving Eta and Nu), the Latin GP has already determined that those were in-script variants.

There are, however, a couple of exceptions. Specifically in the cases of Latin Small Letter I and of Latin Small Letter O with Dot Above, each paired with either an Acute Accent or a Grave Accent:

|  |  |
| --- | --- |
| i U+0069Latin Small Letter I | ì U+00ECLatin Small Letter I with Grave Accent |
| i U+0069Latin Small Letter I | í U+00EDLatin Small Letter I with Acute Accent |
| ȯ U+022F Latin Small Letter O with Dot Above | ò U+00F2Latin Small Letter O with Grave Accent |
| ȯ U+022FLatin Small Letter O with Dot Above | ó U+00F3Latin Small Letter O with Acute Accent |

We are still in the process of evaluating these cases. A response is expected within the next week or so.

Regards

Latin GP

P.S. The Latin repertoire includes some 220 glyphs. In contrast, the Unicode Table includes many times that number of glyphs derived from the Latin alphabet. It occurs to us that the Greek GP might find it useful to know what glyphs are actually in our repertoire. So we have listed them below.

Note that, while the Unicode Table includes large numbers of pre-composed symbols (i.e. a letter plus diacritics), we do have a few cases where it was necessary to use Combining Diacritics (Unicodes 0300 - 036F) to create the necessary glyph.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Unicode** | **Glyph** | **Unicode name** | **Languages using the code point (EGIDS)** | **Reference supporting inclusion (URL etc.)** |
|  | **0061** | **a** | LATIN SMALL LETTER A | Basic Latin | [0] |
|  | **0061 + 0331** | **a̱** | LATIN SMALL LETTER A + COMBINING MACRON BELOW | Nuer (4) | [146], [129] |
|  | **00E0** | **à** | LATIN SMALL LETTER A WITH GRAVE | Italian (1)Galician (2)Wolof (4) | [130], [131], [106], [132] |
|  | **00E1** | **á** | LATIN SMALL LETTER A WITH ACUTE | Spanish (1)Czech (1)Icelandic (1)Faroese (2)Kirundi (1)Chuukese (2)Galician (2)Lule Sámi (2)Northern Sámi (2) | [100], [101], [102], [103], [104], [105], [106], [107], [108] |
|  | **00E2** | **â** | LATIN SMALL LETTER A WITH CIRCUMFLEX | Vietnamese (1)Romanian (1)Skolt Sami (2)Kirundi (1)French (1)Galician (2)West Frisian (2)Friulian (4)Xavante (4) | [109], [110], [113], [104], [114], [106], [115], [116], [117] |
|  | **00E3** | **ã** | LATIN SMALL LETTER A WITH TILDE | Umbundu (3)Guarani (1)Nauruan (3)Khoekhoe (4) | [141], [142], [143], [144], [145] |
|  | **00E4** | **ä** | LATIN SMALL LETTER A WITH DIAERESIS | German (1)Finnish (1)Turkmen (1)Estonian (1)Swedish (1)Lule Sámi (2)Yapese (2)Dinka (4)Kaqchikel (4)Bashkir (4)Alsatian (5)Nuer (4) | [119], [120], [121], [122], [123], [107], [124], [125], [126], [127], [128], [129] |
|  | **00E5** | **å** | LATIN SMALL LETTER A WITH RING ABOVE | Danish (1)Finnish (1) Chamorro (1)Swedish (1)Lule Sámi (2) | [139], [120], [140], [123], [107] |
|  | **00E6** | **æ** | LATIN SMALL LETTER AE | Danish (1)Icelandic (1)Faroese (2) | [139], [102], [103] |
|  | **0101** | **ā** | LATIN SMALL LETTER A WITH MACRON | Latvian (1)Tongan (1)Hawaiian (2)Marshallese(1) | [133], [134], [135], [136] |
|  | **0103** | **ă** | LATIN SMALL LETTER A WITH BREVE | Vietnamese (1)Romanian (1) | [109], [110] |
|  | **0105** | **ą** | LATIN SMALL LETTER A WITH OGONEK | Polish (1)Lithuanian (1) | [137], [138] |
|  | **01CE** | **ǎ** | LATIN SMALL LETTER A WITH CARON | Kirundi (1) | [104] [**https://www.dropbox.com/s/ptfclojxkmbceyf/Kirundi%20and%20its%20tonal%20diacritics.docx**](https://www.dropbox.com/s/ptfclojxkmbceyf/Kirundi%20and%20its%20tonal%20diacritics.docx)**Jean Paul Nkurunziza (personal communication)** |
|  | **1EA1** | **ạ** | LATIN SMALL LETTER A WITH DOT BELOW | Vietnamese (1) | [109] |
|  | **1EA3** | **ả** | LATIN SMALL LETTER A WITH HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EA5** | **ấ** | LATIN SMALL LETTER A WITH CIRCUMFLEX AND ACUTE | Vietnamese (1) | [109] |
|  | **1EA7** | **ầ** | LATIN SMALL LETTER A WITH CIRCUMFLEX AND GRAVE | Vietnamese (1) | [109] |
|  | **1EA9** | **ẩ** | LATIN SMALL LETTER A WITH CIRCUMFLEX AND HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EAB** | **ẫ** | LATIN SMALL LETTER A WITH CIRCUMFLEX AND TILDE | Vietnamese (1) | [109] |
|  | **1EAD** | **ậ** | LATIN SMALL LETTER A WITH CIRCUMFLEXAND DOT BELOW | Vietnamese (1) | [109] |
|  | **1EAF** | **ắ** | LATIN SMALL LETTER A WITH BREVE AND ACUTE | Vietnamese (1) | [109] |
|  | **1EB1** | **ằ** | LATIN SMALL LETTER A WITH BREVE AND GRAVE | Vietnamese (1) | [109] |
|  | **1EB3** | **ẳ** | LATIN SMALL LETTER A WITH BREVE AND HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EB5** | **ẵ** | LATIN SMALL LETTER A WITH BREVE AND TILDE | Vietnamese (1) | [109] |
|  | **1EB7** | **ặ** | LATIN SMALL LETTER A WITH BREVE AND DOT BELOW | Vietnamese (1) | [109] |
|  | **0062** | **b** | LATIN SMALL LETTER B | Basic Latin | [0] |
|  | **0253** | **ɓ** | LATIN SMALL LETTER B WITH HOOK | Hausa (2)Dagaare - Burkina Faso (4) Pulaar, (3) | [147], [148], [250] |
|  | **0063** | **c** | LATIN SMALL LETTER C | Basic Latin | [0] |
|  | **00E7** | **ç** | LATIN SMALL LETTER C WITH CEDILLA | Turkish (1)Turkmen (1)Kurdish (2)French (1)Azerbaijani (1)Basque (1)Galician (2)Friulian (4)Bashkir (4) | [157], [121], [158], [114], [159], [160], [161], [106], [116], [127] |
|  | **0107** | **ć** | LATIN SMALL LETTER C WITH ACUTE | Croatian (1)Serbian (1)Polish (1) | [150], [151], [152] |
|  | **0109** | **ĉ** | LATIN SMALL LETTER C WITH CIRCUMFLEX | Esperanto (3) | [[255]](http://www.omniglot.com/writing/esperanto.htm) |
|  | **010B** | **ċ** | LATIN SMALL LETTER C WITH DOT ABOVE | Maltese (1) | [163] |
|  | **010D** | **č** | LATIN SMALL LETTER C WITH CARON | Croatian (1)Serbian (1)Latvian (1)Slovak (1)Northern Sámi (2)Lithuanian (1) | [150], [151], [133], [153], [108], [154]  |
|  | **0064** | **d** | LATIN SMALL LETTER D | Basic Latin  | [0] |
|  | **00F0** | **ð** | LATIN SMALL LETTER ETH | Faroese (2)Icelandic (1) | [103], [102] |
|  | **010F** | **ď** | LATIN SMALL LETTER D WITH CARON | Czech (1)Slovak (1) | [101], [153] |
|  | **0111** | **đ** | LATIN SMALL LETTER D WITH STROKE | Croatian (1)Serbian (1)Vietnamese (1)Northern SámiBrahui (5) | [150], [151], [109], [108], [168] |
|  | **0256** | **ɖ** | LATIN SMALL LETTER D WITH TAIL | Fon (3)Ewe (3) | [169], [170] |
|  | **0257** | **ɗ** | LATIN SMALL LETTER D WITH HOOK | Hausa (2)Pulaar (3) | [147], [166], [250] |
|  | **1E13** | **ḓ** | LATIN SMALL LETTER D WITH CIRCUMFLEX BELOW | Venda (1) | [164], [257] |
|  | **0065** | **e** | LATIN SMALL LETTER E | Basic Latin | [0] |
|  | **0065 + 0331** | **e̱** | LATIN SMALL LETTER E + COMBINING MACRON BELOW | Nuer (4) | [146] |
|  | **00E8** | **è** | LATIN SMALL LETTER E WITH GRAVE  | French (1)Italian (1) Afrikaans (1)Kirundi (1)Haitian Creole (1) French (1) | [114], [130], [175], [104], [182], [183] |
|  | **00E9** | **é** | LATIN SMALL LETTER E WITH ACUTE | French (1)Italian (1) Spanish (1)Czech (1)Icelandic (1)Kirundi (1)Chuukese (2)Galician (2)Wolof (4)XAVANTE (4)West Frisian (2) | [114], [130], [100], [101], [102], [104], [105], [106], [132], [117], [115] |
|  | **00EA** | **ê** | LATIN SMALL LETTER E WITH CIRCUMFLEX | French (1)Tswana (1)Afrikaans (1)Vietnamese (1)Kurdish (2)Kirundi (1)West Frisian (2)Friulian (4) | [114], [173], [174], [175], [109], [158], [104], [115], [116] |
|  | **00EB** | **ë** | LATIN SMALL LETTER E WITH DIAERESIS | Afrikaans (1)Kirundi (1)Albanian (1)French (1)Chuukese (2)Uyghur (2)Yapese (2)Wolof (4)Drehu (4)Kaqchikel (4)West Frisian (2)Nuer (4) | [175], [104], [176], [177], [114], [176], [177], [114], [178], [179], [124], [132], [180], [126], [115], [129] |
|  | **0113** | **ē** | LATIN SMALL LETTER E WITH MACRON | Latvian (1)Hawaiian (2)Tongan (1)Minangkabau (5) | [133], [135], [134], [184] |
|  | **0117** | **ė** | LATIN SMALL LETTER E WITH DOT ABOVE | Lithuanian (1) | [138], [154] |
|  | **0119** | **ę** | LATIN SMALL LETTER E WITH OGONEK | Polish (1)Palauan (2)Lithuanian (1) | [152], [185], [138], [154] |
|  | **011B** | **ě** | LATIN SMALL LETTER E WITH CARON | Czech (1)Kirundi (1)Sorbian (4) | [101], [104], [172] |
|  | **01DD** | **ǝ** | LATIN SMALL LETTER TURNED E | Kanuri (3) | [240] |
|  | **0259** | **ə** | LATIN SMALL LETTER SCHWA | Azeri, Azerbaijani (1)Ewondo (3)Ewe (3)Bugis (3) | [159], [190], [170], [241] |
|  | **025B** | **ɛ** | LATIN SMALL LETTER OPEN E | Dagaare - Burkina Faso (4)Lingala (2)Akan (3)Ewondo (3)Dagbani (Dagomba) (4)Fon (3)Mossi (3)Ga (4)Ewe (3)Duala (3)Bambara (4)Nuer (4) | [148], [236], [237], [190], [189], [169], [212], [238], [193], [170], [194], [199], [129] |
|  | **025B + 0308** | **ɛ̈** | LATIN SMALL LETTER OPEN E + COMBINING DIAERESIS | Nuer (4)Dinka (4) | [129], [146], [239], [125] |
|  | **025B + 0331** | **ɛ̱** | LATIN SMALL LETTER OPEN E + COMBINING MACRON BELOW | Nuer (4) | [129], [146], [239] |
|  | **025B + 0331 + 0308** | **ɛ̱̈** | LATIN SMALL LETTER OPEN E + COMBINING MACRON BELOW + COMBINING DIAERESIS  | Nuer (4) | [146], [239] |
|  | **1EB9** | **ẹ** | LATIN SMALL LETTER E WITH DOT BELOW | Yoruba (2) | [181] |
|  | **1EB9 + 0300** | **ẹ̀** | LATIN SMALL LETTER E WITH DOT BELOW + COMBINING GRAVE ACCENT | Yoruba (2) | [254] |
|  | **1EB9 + 0301** | **ẹ́** | LATIN SMALL LETTER E WITH DOT BELOW + COMBINING ACUTE ACCENT | Yoruba (2) | [254] |
|  | **1EBB** | **ẻ** | LATIN SMALL LETTER E WITH HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EBD** | **ẽ** | LATIN SMALL LETTER E WITH TILDE | Umbundu (3)Guarani (1)Cubeo (3)Xavante (4) | [141], [142], [143], [186], [187], [117] |
|  | **1EBF** | **ế** | LATIN SMALL LETTER E WITH CIRCUMFLEX AND ACUTE | Vietnamese (1) | [109] |
|  | **1EC1** | **ề** | LATIN SMALL LETTER E WITH CIRCUMFLEX AND GRAVE | Vietnamese (1) | [109] |
|  | **1EC3** | **ể** | LATIN SMALL LETTER E WITH CIRCUMFLEX AND HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EC5** | **ễ** | LATIN SMALL LETTER E WITH CIRCUMFLEX AND TILDE | Vietnamese (1) | [109] |
|  | **1EC7** | **ệ** | LATIN SMALL LETTER E WITH CIRCUMFLEXAND DOT BELOW | Vietnamese (1) | [109] |
|  | **0066** | **f** | LATIN SMALL LETTER F | Basic Latin | [0] |
|  | **0192** | **ƒ** | LATIN SMALL LETTER F WITH HOOK | Ewe (3) | [170] |
|  | **0067** | **g** | LATIN SMALL LETTER G | Basic Latin | [0] |
|  | **0067 + 0303** | **g̃** | LATIN SMALL LETTER G + COMBINING TILDE | Guarani (1) | [142], [143] |
|  | **0067 + 0304** | **ḡ** | LATIN SMALL LETTER G + COMBINING MACRON | Raga (Hano) (3) | [200] |
|  | **011D** | **ĝ** | LATIN SMALL LETTER G WITH CIRCUMFLEX | Esperanto (3) | [[255]](http://www.omniglot.com/writing/esperanto.htm) |
|  | **011F** | **ğ** | LATIN SMALL LETTER G WITH BREVE | Turkish (1)Tatar (2) Azeri (1)Bashkir (4)Zaza (5) | [157], [201], [159], [127], [202] |
|  | **0121** | **ġ** | LATIN SMALL LETTER G WITH DOT ABOVE | Maltese (1) | [163] |
|  | **0123** | **ģ** | LATIN SMALL LETTER G WITH CEDILLA | Latvian (1)Brahui (5) | [133], [168] |
|  | **01E7** | **ǧ** | LATIN SMALL LETTER G WITH CARON | Skolt Sami (2) | [113] |
|  | **0263** | **ɣ** | LATIN SMALL LETTER GAMMA | Dagbani (Dagomba) (4)Nuer (4)Dinka (4)Ewe (3)Nuer (4) | [189], [146], [125], [170], [129] |
|  | **0068** | **h** | LATIN SMALL LETTER H | Basic Latin | [0] |
|  | **0125** | **ĥ** | LATIN SMALL LETTER H WITH CIRCUMFLEX | Esperanto (3) | [255] |
|  | **0127** | **ħ** | LATIN SMALL LETTER H WITH STROKE | Maltese (1) | [163] |
|  | **0069** | **i** | LATIN SMALL LETTER I | Basic Latin | [0] |
|  | **0069 + 0331** | **i̲** | LATIN SMALL LETTER I + COMBINING MACRON BELOW | Nuer (4) | [146] |
|  | **00EC** | **ì** | LATIN SMALL LETTER I WITH GRAVE | Italian (1)Kirundi (1) | [130], [206], [208] |
|  | **00ED** | **í** | LATIN SMALL LETTER I WITH ACUTE | Spanish (1)Czech (1) Icelandic (1) Faroese (2) Kirundi (1)Galician (2)Bashkir(4) | [100], [101], [102], [103], [104], [106], [127] |
|  | **00EE** | **î** | LATIN SMALL LETTER I WITH CIRCUMFLEX | Afrikaans (1)Romanian (1)Kurdish (2) Kirundi (1) French (1)Friulian (4) | [175], [110], [158], [104], [114], [116] |
|  | **00EF** | **ï** | LATIN SMALL LETTER I WITH DIAERESIS | Afrikaans (1)French (1)Kaqchikel (4)Dinka (4)West Frisian (2) | [175], [114], [126], [125], [115] |
|  | **0129** | **ĩ** | LATIN SMALL LETTER I WITH TILDE | Guarani (1)Cubeo (3)Khoekhoe (4)Kikuyu ( 5) | [142], [143], [186], [145], [209] |
|  | **012B** | **ī** | LATIN SMALL LETTER I WITH MACRON | Latvian (1)Lithuanian (1) Hawaiian (2)Tongan (1) | [133], [138], [135], [134] |
|  | **012F** | **į** | LATIN SMALL LETTER I WITH OGONEK | Lithuanian (1) | [154] |
|  | **0131** | **ı** | LATIN SMALL LETTER I DOTLESS | Turkish (1)Tatar (2) Azeri (1) | [157], [203], [201], [159] |
|  | **0135** | **ĵ** | LATIN SMALL LETTER J WITH CEDILLA | Esperanto (3) | [255] |
|  | **01D0** | **ǐ** | LATIN SMALL LETTER I WITH CARON | Kirundi (1) | [104] |
|  | **0268** | **ɨ** | LATIN SMALL LETTER I WITH STROKE | Cubeo (3)Dagbani (Dagomba) (4)HIxkaryána (4)Maasai (5) | [186], [189], [210], [211] |
|  | **0268 + 0303** | **ɨ̃** | LATIN SMALL LETTER I WITH STROKE + COMBINING TILDE | Cubeo (3) | [186] |
|  | **1EC9** | **ỉ** | LATIN SMALL LETTER I WITH HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1ECB** | **ị** | LATIN SMALL LETTER I WITH DOT BELOW | Igbo (2) | [205] |
|  | **006A** | **j** | LATIN SMALL LETTER J | Basic Latin | [0] |
|  | **0269** | **ɩ** | LATIN SMALL LETTER IOTA | Dagaare - Burkina Faso (4)Mossi (3) | [148], [212] |
|  | **006B** | **k** | LATIN SMALL LETTER K | Basic Latin | [0] |
|  | **0137** | **ķ** | LATIN SMALL LETTER K WITH CEDILLA | Latvian (1) | [133] |
|  | **0199** | **ƙ** | LATIN SMALL LETTER K WITH HOOK | Hausa (2) | [147] |
|  | **01E9** | **ǩ** | LATIN SMALL LETTER K WITH CARON | Skolt Sami (2) | [113] |
|  | **006C** | **l** | LATIN SMALL LETTER L | Basic Latin | [0] |
|  | **013A** | **ĺ** | LATIN SMALL LETTER L WITH ACUTE | Slovak (1) | [153] |
|  | **013C** | **ļ** | LATIN SMALL LETTER L WITH CEDILLA | Latvian (1)Marshallese (1)Brahui (5) | [133], [213], [214], [168] |
|  | **013E** | **ľ** | LATIN SMALL LETTER L WITH CARON | Slovak (1) | [153] |
|  | **0142** | **ł** | LATIN SMALL LETTER L WITH STROKE | Polish (1) | [152] |
|  | **1E37** | **ḷ** | LATIN SMALL LETTER L WITH DOT BELOW | Marshallese (1) | [213], [214], [215], [216] |
|  | **1E3D** | **ḽ** | LATIN SMALL LETTER L WITH CIRCUMFLEX BELOW | Venda (1) | [164], [257] |
|  | **006D** | **m** | LATIN SMALL LETTER M | Basic Latin | [0] |
|  | **006D + 0327** | **m̦** | LATIN SMALL LETTER M + COMBINING CEDILLA | Marshallese (1) | [213], [136], [214] |
|  | **1E43** | **ṃ** | LATIN SMALL LETTER M WITH DOT BELOW | Marshallese (1) | [213], [136], [215], [216] |
|  | **006E** | **n** | LATIN SMALL LETTER N | Basic Latin | [0] |
|  | **006E + 0304** | **n̄** | LATIN SMALL LETTER N + COMBINING MACRON | Raga (Hano) (3)Marshallese (1) | [200], [213], [136] |
|  | **006E + 0308** | **n̈** | LATIN SMALL LETTER N + COMBINING DIAERESIS |  Malagasy(1) | [230] |
|  | **00F1** | **ñ** | LATIN SMALL LETTER N WITH TILDE | Spanish (1)Pulaar (3)Chamorro (1)Filipino (1)Guarani (1)Chavacano (4)Basque (1)Galician (2)Iloco (3)Quechua (3)Cape Verdean Creole (4)Waray-Waray (3)Wolof (4)Nauruan (3)Lozi (4)Bashkir (4)Marshallese (1)Mandinka (5)Igbo(2) | [221], [250] [222], [142], [143], [223], [160], [106], [224], [225], [226], [227], [228], [132], [144], [229], [127], [136], [197], [205] |
|  | **0144** | **ń** | LATIN SMALL LETTER N WITH ACUTE | Polish (1)Lule Sámi (2)Sorbian (4)Brahui (5) | [152], [107], [172], [168] |
|  | **0146** | **ņ** | LATIN SMALL LETTER N WITH CEDILLA | Latvian (1)Marshallese (1) | [133], [136] |
|  | **0148** | **ň** | LATIN SMALL LETTER N WITH CARON | Turkmen (1)Czech (1)Slovak (1) | [121], [101], [153] |
|  | **014B** | **ŋ** | LATIN SMALL LETTER ENG | Inari Sami (2)Dagaare - Burkina Faso (4)Dagbani (Dagomba) (4)Northern Sami (2)Ewondo (3)Luganda (3)Wolof (4)Adzera (4)Nuer (4)Ga (4)Dinka (4)Duala (3)Ewe (3)Soga (5) Alur (5) Mandinka (5) Acholi (5)Bambara (4)Nuer (4) | [188], [148], [189], [108], [190], [191], [132], [192], [146], [193], [125], [194], [170], [195], [196], [197], [198], [199], [129] |
|  | **0272** | **ɲ** | LATIN SMALL LETTER N WITH LEFT HOOK | Susu (4)Zarma (4)Bambara (4) | [218], [219], [199] |
|  | **1E45** | **ṅ** | LATIN SMALL LETTER N WITH DOT ABOVE | Venda (1) | [164], [257] |
|  | **1E47** | **ṇ** | LATIN SMALL LETTER N WITH DOT BELOW | Marshallese (1) | [136], [215], [216] |
|  | **1E49** | **ṉ** | LATIN SMALL LETTER N WITH LINE BELOW | Pitjantjatjara (4) | [220] |
|  | **1E4B** | **ṋ** | LATIN SMALL LETTER N WITH CIRCUMFLEX BELOW | Venda (1) | [164], [257] |
|  | **006F** | **o** | LATIN SMALL LETTER O | Basic Latin | [0] |
|  | **006F + 0327** | **o̧** | LATIN SMALL LETTER O + COMBINING CEDILLA | Marshallese (1) | [136] |
|  | **006F + 0331** | **o̱** | LATIN SMALL LETTER O + COMBINING MACRON BELOW | Nuer (4) | [146], [129] |
|  | **00F2** | **ò** | LATIN SMALL LETTER O WITH GRAVE | Italian (1)Haitian Creole (1) | [130], [182], [183] |
|  | **00F3** | **ó** | LATIN SMALL LETTER O WITH ACUTE | Spanish (1)Polish (1)Czech (1)Icelandic (1)Kirundi (1)Chuukese (2)Galician (2)Wolof (4) | [100], [152], [101], [102], [104], [105], [106], [132] |
|  | **00F4** | **ô** | LATIN SMALL LETTER O WITH CIRCUMFLEX | Tswana (1)Afrikaans (1)Vietnamese (1)Kirundi (1)French (1)Northern Sotho (1)West Frisian (2)Galician (2)Friulian (4)Xavante(4) | [173], [174], [175], [109], [104], [114], [230], [115], [106], [116], [117] |
|  | **00F5** | **õ** | LATIN SMALL LETTER O WITH TILDE | Estonian (1)Skolt Sami (2)Umbundu (3)Guarani (1)Nauruan (3)Xavante (4)Khoekhoe (4) | [122], [113], [141], [142], [143], [144], [117], [235] |
|  | **00F6** | **ö** | LATIN SMALL LETTER O WITH DIAERESIS | German (1)Finnish (1)Afrikaans (1)Turkish (1)Swedish (1)Uygur (2)Yapese (2)Drehu (4)Kaqchikel (4)Dinka (4)Bashkir (4)Low German (5)Chechen (2) 1992 VersionWest Frisian (2)Nuer (4) | [119], [120], [175], [157], [123], [179], [124], [180], [126], [125], [127], [231], [232], [115], [129] |
|  | **00F8** | **ø** | LATIN SMALL LETTER O WITH STROKE | Danish (1)Faroese (2) | [139], [103] |
|  | **014D** | **ō** | LATIN SMALL LETTER O WITH MACRON | Hawaiian (2)Marshallese (1)Tongan (1) | [135], [136], [134] |
|  | **0151** | **ő** | LATIN SMALL LETTER O WITH DOUBLE ACUTE | Hungarian (1) | [233], [234] |
|  | **0153** | **œ** | LATIN SMALL LIGATURE OE | French (1) | [114], [253] |
|  | **01A1** | **ơ** | LATIN SMALL LETTER O WITH HORN | Vietnamese (1) | [109] |
|  | **01D2** | **ǒ** | LATIN SMALL LETTER O WITH CARON | Kirundi (1) | [104] |
|  | **0254** | **ɔ** | LATIN SMALL LETTER OPEN O | Dagaare - Burkina Faso (4)Dagbani (Dagomba) (4)Lingala (2)Akan (3)Ewondo (3)Fon (3)Nuer (4)Ga (4)Duala (3)Ewe (3)Nuer (4) | [148], [189], [236], [237], [190], [169], [146], [193], [194], [170], [129] |
|  | **0254 + 0308** | **ɔ̈** | LATIN SMALL LETTER OPEN O + COMBINING DIAERESIS | Dinka (4) | [125] |
|  | **0254 + 0331** | **ɔ̱** | LATIN SMALL LETTER OPEN O + COMBINING MACRON BELOW | Nuer (4) | [129], [146] |
|  | **1ECD** | **ọ** | LATIN SMALL LETTER O WITH DOT BELOW | Igbo (2)Yoruba (2)Marshallese (1) | [204], [205], [181], [136], [215], [216] |
|  | **1ECD + 0300** | **ọ̀** | LATIN SMALL LETTER O WITH DOT BELOW + COMBINING GRAVE ACCENT | Yoruba (2) | [254] |
|  | **1ECD + 0301** | **ọ́** | LATIN SMALL LETTER O WITH DOT BELOW + COMBINING ACUTE ACCENT | Yoruba (2) | [254] |
|  | **1ECF** | **ỏ** | LATIN SMALL LETTER O WITH HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1ED1** | **ố** | LATIN SMALL LETTER O WITH CIRCUMFLEX AND ACUTE | Vietnamese (1) | [109] |
|  | **1ED3** | **ồ** | LATIN SMALL LETTER O WITH CIRCUMFLEX AND GRAVE | Vietnamese (1) | [109] |
|  | **1ED5** | **ổ** | LATIN SMALL LETTER O WITH CIRCUMFLEX AND HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1ED7** | **ỗ** | LATIN SMALL LETTER O WITH CIRCUMFLEX AND TILDE | Vietnamese (1) | [109] |
|  | **1ED9** | **ộ** | LATIN SMALL LETTER O WITH CIRCUMFLEX AND DOT BELOW | Vietnamese (1) | [109] |
|  | **1EDB** | **ớ** | LATIN SMALL LETTER O WITH HORN AND ACUTE | Vietnamese (1) | [109] |
|  | **1EDD** | **ờ** | LATIN SMALL LETTER O WITH HORN AND GRAVE | Vietnamese (1) | [109] |
|  | **1EDF** | **ở** | LATIN SMALL LETTER O WITH HORN AND HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EE1** | **ỡ** | LATIN SMALL LETTER O WITH HORN AND TILDE | Vietnamese (1) | [109] |
|  | **1EE3** | **ợ** | LATIN SMALL LETTER O WITH HORN AND DOT BELOW | Vietnamese (1) | [109] |
|  | **0070** | **p** | LATIN SMALL LETTER P  | Basic Latin | [0] |
|  | **00FE** | **þ** | LATIN SMALL LETTER THORN | Icelandic (1) | [102] |
|  | **0071** | **q** | LATIN SMALL LETTER Q | Basic Latin | [0] |
|  | **0072** | **r** | LATIN SMALL LETTER R | Basic Latin | [0] |
|  | **0072 + 0303** | **r̃** | LATIN SMALL LETTER R + COMBINING TILDE | Hausa (2) | [147] |
|  | **0155** | **ŕ** | LATIN SMALL LETTER R WITH ACUTE | Slovak (1)Brahui (5) | [153], [168] |
|  | **0159** | **ř** | LATIN SMALL LETTER R WITH CARON | Czech (1)Sorbian (4) | [101], [172] |
|  | **024D** | **ɍ** | LATIN SMALL LETTER R WITH STROKE | Kanuri (3) | [240] |
|  | **0073** | **s** | LATIN SMALL LETTER S | Basic Latin | [0] |
|  | **00DF** | **ß** | LATIN SMALL LETTER SHARP S | German (1) | [119] |
|  | **015B** | **ś** | LATIN SMALL LETTER S WITH ACUTE | Polish (1) Montenegrin (1) | [152], [258]  |
|  | **015D** | **ŝ** | LATIN SMALL LETTER S WITH CIRCUMFLEX | Esperanto (3) | [255] |
|  | **015F** | **ş** | LATIN SMALL LETTER S WITH CEDILLA | Turkish (1)Turkmen (1)Kurdish (2)Tatar (2)Azeri (1)Bashkir (4)Brahui (5)Zaza (5) | [157], [121], [158], [201], [159], [127], [168], [202] |
|  | **0161** | **š** | LATIN SMALL LETTER S WITH CARON | Tswana (1)Croatian (1)Serbian (1)Latvian (1)Northern Sotho (1)Northern Sami (2)Lithuanian (1) | [174], [150], [151], [133], [230], [108], [154] |
|  | **0219** | **ș** | LATIN SMALL LETTER S WITH COMMA BELOW | Romanian (1) | [110] |
|  | **1E63** | **ṣ** | LATIN SMALL LETTER S WITH DOT BELOW | Yoruba (2) | [181] |
|  | **0074** | **t** | LATIN SMALL LETTER T | Basic Latin | [0] |
|  | **0165** | **ť** | LATIN SMALL LETTER T WITH CARON | Czech (1)Slovak (1) | [101], [153] |
|  | **0167** | **ŧ** | LATIN SMALL LETTER T WITH STROKE | Northern Sami (2)Brahui (5) | [108], [168] |
|  | **021B** | **ț** | LATIN SMALL LETTER T WITH COMMA BELOW | Romanian (1) | [110] |
|  | **1E6D** | **ṭ** | LATIN SMALL LETTER T WITH DOT BELOW | Mizo (4) | [242] |
|  | **1E71** | **ṱ** | LATIN SMALL LETTER T WITH CIRCUMFLEX BELOW | Venda (1) | [164]**, [257]** |
|  | **0075** | **u** | LATIN SMALL LETTER U | Basic Latin | [0] |
|  | **00F9** | **ù** | LATIN SMALL LETTER U WITH GRAVE | Italian (1)Papiamento (1) | [130], [206], [245], [246] |
|  | **00FA** | **ú** | LATIN SMALL LETTER U WITH ACUTE | Spanish (1)Czech (1)Icelandic (1)Faroese (2)Kirundi (1)Chuukese (2)West Frisian (2)Galician (2) | [100], [101], [102], [103], [104], [105], [115], [106] |
|  | **00FB** | **û** | LATIN SMALL LETTER U WITH CIRCUMFLEX | Afrikaans (1) Kurdish (2)Kirundi (1)French (1)Miskito (2)West Frisian (2)Friulian (4)Zazaki (4) | [175], [158], [104], [114], [243], [115], [116], [244] |
|  | **00FC** | **ü** | LATIN SMALL LETTER U WITH DIAERESIS | German (1)Spanish (1)Afrikaans (1)Turkish (1)Swedish (1)French (1)Azeri (1)Basque (1)Galician (2)Uygur (2)Kaqchikel (4)Bashkir (4) | [119], [100], [175], [157], [123], [114], [159], [161], [106], [179], [126], [127], [231] |
|  | **0169** | **ũ** | LATIN SMALL LETTER U WITH TILDE | Umbundu (3)Guarani (1)Nauruan (3)Khoekhoe (4)Kikuyu (5) | [141], [142], [143], [144], [145], [209] |
|  | **016B** | **ū** | LATIN SMALL LETTER U WITH MACRON | Latvian (1)Hawaiian (2)Lithuanian (1)Marshallese (1)Tongan (1) | [133], [135], [138], [154], [136], [134] |
|  | **016D** | **ŭ** | LATIN SMALL LETTER U WITH BREVE | Esperanto (3) | [255] |
|  | **016F** | **ů** | LATIN SMALL LETTER U WITH RING ABOVE | Czech (1) | [101] |
|  | **0171** | **ű** | LATIN SMALL LETTER U WITH DOUBLE ACUTE | Hungarian (1) | [233], [234] |
|  | **0173** | **ų** | LATIN SMALL LETTER U WITH OGONEK | Lithuanian (1) | [154], [138] |
|  | **01B0** | **ư** | LATIN SMALL LETTER U WITH HORN | Vietnamese (1) | [109] |
|  | **01D4** | **ǔ** | LATIN SMALL LETTER U WITH CARON | Kirundi (1) | [104] |
|  | **0289** | **ʉ** | LATIN SMALL LETTER U BAR | Cubeo (3)Maasai (5) | [186], [187], [211] |
|  | **0289 + 0303** | **ʉ̃** | LATIN SMALL LETTER U BAR + COMBINING TILDE | Cubeo (3) | [186], [187] |
|  | **1EE5** | **ụ** | LATIN SMALL LETTER U WITH DOT BELOW | Vietnamese (1)Igbo (2) | [109],[204], [205] |
|  | **1EE7** | **ủ** | LATIN SMALL LETTER U WITH HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EE9** | **ứ** | LATIN SMALL LETTER U WITH HORN AND ACUTE | Vietnamese (1) | [109] |
|  | **1EEB** | **ừ** | LATIN SMALL LETTER U WITH HORN AND GRAVE | Vietnamese (1) | [109] |
|  | **1EED** | **ử** | LATIN SMALL LETTER U WITH HORN AND HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EEF** | **ữ** | LATIN SMALL LETTER U WITH HORN AND TILDE | Vietnamese (1) | [109] |
|  | **1EF1** | **ự** | LATIN SMALL LETTER U WITH HORN AND DOT BELOW | Vietnamese (1) | [109] |
|  | **0076** | **v** | LATIN SMALL LETTER V | Basic Latin | [0] |
|  | **028B** | **ʋ** | LATIN SMALL LETTER V WITH HOOK | Dagaare - Burkina Faso (4)Mossi (3)Ewe (3) | [148], [212], [238], [170] |
|  | **0077** | **w** | LATIN SMALL LETTER W | Basic Latin | [0] |
|  | **0175** | **ŵ** | LATIN SMALL LETTER W WITH CIRCUMFLEX | Chichewa (3) | [247] |
|  | **0078** | **x** | LATIN SMALL LETTER X | Basic Latin | [0] |
|  | **1E8D** | **ẍ** | LATIN SMALL LETTER X WITH DIAERESIS | Mam (4) | [248], [249] |
|  | **0079** | **y** | LATIN SMALL LETTER Y | Basic Latin | [0] |
|  | **00FD** | **ý** | LATIN SMALL LETTER Y WITH ACUTE | Turkmen (1)Czech (1)Icelandic (1)Faroese (2)Guarani (1) | [121], [101], [102], [103], [142], [143] |
|  | **00FF** | **ÿ** | LATIN SMALL LETTER Y WITH DIAERESIS | French (1) | [114], [253], [257] |
|  | **0177** | **ŷ** | LATIN SMALL LETTER Y WITH CIRCUMFLEX | Welsh (2) | [256] |
|  | **01B4** | **ƴ** | LATIN SMALL LETTER Y WITH HOOK | Dagaare - Burkina Faso (4) | [148], [251], [149] |
|  | **1EF3** | **ỳ** | LATIN SMALL LETTER Y WITH GRAVE | Vietnamese (1) | [109] |
|  | **1EF5** | **ỵ** | LATIN SMALL LETTER Y WITH DOT BELOW | Vietnamese (1) | [109] |
|  | **1EF7** | **ỷ** | LATIN SMALL LETTER Y WITH HOOK ABOVE | Vietnamese (1) | [109] |
|  | **1EF9** | **ỹ** | LATIN SMALL LETTER Y WITH TILDE | Vietnamese (1) Guarani (1) | [109] [142] |
|  | **007A** | **z** | LATIN SMALL LETTER Z | Basic Latin | [0] |
|  | **017A** | **ź** | LATIN SMALL LETTER Z WITH ACUTE | Polish (1)Brahui (5)Sorbian (4)Montenegrin(1) | [152], [252], [168], [172], [258]  |
|  | **017C** | **ż** | LATIN SMALL LETTER Z WITH DOT ABOVE | Polish (1)Maltese (1) | [152], [163] |
|  | **017E** | **ž** | LATIN SMALL LETTER Z WITH CARON | Lithuanian (1)Croatian (1)Serbian (1)Turkmen (1)Latvian (1)Slovak (1)Northern Sami (2)Chechen (2) 1925 Version | [154], [150], [151], [121], [133], [153], [108], [232] |
|  | **01EF** | **ǯ** | LATIN SMALL LETTER EZH WITH CARON | Skolt Sami (2) | [113] |
|  | **0292** | **ʒ** | LATIN SMALL LETTER EZH | Skolt Sami (2)Dagbani (Dagomba) (4) | [113], [189] |