

Re: Script_Extension values for some combining marks
 From: Mark Davis
 Date: 2017-...

The following Script_Extension property values are proposed for addition in Unicode 11.0.

```
# Script_Extensions=Cyrl
030F ;          Cyrl # COMBINING DOUBLE GRAVE ACCENT

# Script_Extensions=Latn
0302..0303 ;    Latn # COMBINING CIRCUMFLEX ACCENT..COMBINING TILDE
0307 ;          Latn # COMBINING DOT ABOVE
0309..030A ;    Latn # COMBINING HOOK ABOVE..COMBINING RING ABOVE
030C..030D ;    Latn # COMBINING CARON..COMBINING VERTICAL LINE ABOVE
0311 ;          Latn # COMBINING INVERTED BREVE
031B ;          Latn # COMBINING HORN
0320 ;          Latn # COMBINING MINUS SIGN BELOW
0323 ;          Latn # COMBINING DOT BELOW
0326..0328 ;    Latn # COMBINING COMMA BELOW..COMBINING OGONEK
032D ;          Latn # COMBINING CIRCUMFLEX ACCENT BELOW
0330..0331 ;    Latn # COMBINING TILDE BELOW..COMBINING MACRON BELOW
0338 ;          Latn # COMBINING LONG SOLIDUS OVERLAY
035F ;          Latn # COMBINING DOUBLE MACRON BELOW
1DC5..1DC7 ;    Latn # COMBINING GRAVE-MACRON..COMBINING ACUTE-MACRON

# Script_Extensions=Grek
0313..0314 ;    Grek # COMBINING COMMA ABOVE..COMBINING REVERSED COMMA ABOVE

# Script_Extensions=Cyrl Latn
0304 ;          Cyrl Latn # COMBINING MACRON
0306 ;          Cyrl Latn # COMBINING BREVE
030B ;          Cyrl Latn # COMBINING DOUBLE ACUTE ACCENT

# Script_Extensions=Cyrl Latn Grek
0300..0301 ;    Cyrl Latn Grek # COMBINING GRAVE ACCENT..COMBINING ACUTE ACCENT
0308 ;          Cyrl Latn Grek # COMBINING DIAERESIS
0335 ;          Cyrl Latn Grek # COMBINING SHORT STROKE OVERLAY
```

Background

The Script_Extensions data supplements the Script value by providing information as to the known usage of characters. Among other uses, Script_Extensions are important in the processing used for spoof detection in UTS #39.

The values for the above characters were obtained in the following way:

The exemplar data in CLDR (which includes a large amount of information donated by SIL), was processed. If a character in a main or aux exemplar set was a combining character, the script of the main exemplar set was added to the Script_Extensions value of that combining character.

