

A note on U+1AB3 COMBINING DOWNWARDS ARROW

Kirk Miller, kirkmiller, gmail.com

2025 February 6

The attestation in L2/11-202 for U+1AB3 COMBINING DOWNWARDS ARROW ◌↓ is a typographic substitute for the fusion of U+30C COMBINING CARON and U+301 COMBINING ACUTE ACCENT. The intended diacritic is used for example in Kajkavian dialectology on the consonants /č/, /šč/ and /ž/. In some sources these letters bear a discrete caron and vertical stroke, <č̣ ẓ̌>, or caron and acute accent, <č̣̇ ẓ̌̇>; in others the two diacritics are overstruck or even fused into a crown shape, <č̣̣̇ ẓ̣̌̇>, as in Figure 1. There is no attested source where an actual arrow was intended.

Semantically as well, this diacritic is seen to be a compound of caron and acute in Figure 2 and Figure 3. Note that in Figure 2 the diacritic is graphically caron + vertical stroke despite being identified with the acute accent; the vertical orientation of the compounded acute accent is a result of its historical origin in italic typeface and may have been an accommodation to legibility.

The Unicode proposal for this character, L2/11-202, has a single illustration (reproduced in Error: Reference source not found), where a downward arrow is indeed substituted for the intended diacritic. The description — *a voiceless sibilant, between š and χ*, where in the German tradition <χ> is a palatal fricative equivalent to IPA <ç> — makes it clear that it is indeed the same diacritic, while the fact that the s is dropped below the baseline shows this token to be a typographic hack.

The convention of compounding the caron and acute for this use traces back at least to Lepsius in the mid 19th century, where the acute accent is set at a nearly vertical angle and nests naturally inside the caron when set in italic typeface, as seen with <ṧ> and <ž̇> in Figure 5.

The question then is how to rectify this error within Unicode. The SEW suggested that a comment be added to the Core Spec, in section 7.9.2 Combining Diacritical Marks Extended, as well as an annotation to U+1AB3, and that U+1AB3 be left as an arrow diacritic.

An advantage of this solution is that if there is ever need for a downward arrow diacritic in the future, the name of U+1AB3 is the most appropriate one for it. Currently the naming of U+1AB3 fits comfortably within the set of arrow diacritics in Unicode:

- ◌↑ U+034E COMBINING UPWARDS ARROW BELOW.
- ◌← U+20EE COMBINING LEFT ARROW BELOW.
- ◌→ U+20EF COMBINING RIGHT ARROW BELOW.
- ◌↔ U+034D COMBINING LEFT RIGHT ARROW BELOW.
- ◌↓ U+1AB3 COMBINING DOWNWARDS ARROW BELOW (L2/24-080).

- ◌ U+20D6 COMBINING LEFT ARROW ABOVE.
- ◌ U+20D7 COMBINING RIGHT ARROW ABOVE.
- ◌ U+20E1 COMBINING LEFT RIGHT ARROW ABOVE.
- ◌ U+1AB3 COMBINING DOWNWARDS ARROW ABOVE.

By now there may be published sources that use U+1AB3 as an arrow because of its name. If people have been using U+1AB3 as an arrow, so that its inclusion in Unicode as an arrow is now justified by use, we would not necessarily realize this because users would not send attestation for a character that already exists. The SEW recommends that typographers not change the glyph for U+1AB3 in their fonts – e.g. Gentium <◌>, Andika <◌>, Noto Serif <◌>, Noto Sans <◌>.

Core Spec

I suggest adding a paragraph like the following to the Core Spec, under section 7.9.2 Combining Diacritical Marks Extended:

Combining downwards arrow. The source glyph was a printer’s hack for a conflated caron and acute; the acute tends to become vertical in this combination due to a history of being set in italic typeface. For Croatian dialectology, use the sequence U+030C COMBINING CARON + 0301 COMBINING ACUTE ACCENT (thus: <◌>) instead of U+1AB3.

Annotation

1AB3 COMBINING DOWNWARDS ARROW ABOVE

- for Croatian dialectology, use the sequence U+030C COMBINING CARON + 0301 COMBINING ACUTE ACCENT

References

- Belović & Blažeka (2015) Dopune rječniku svetog đurđa (rječniku ludbreške podravine), *Jezičnica Kajkaviana, KAJ – časopis za književnost, umjetnost, kulturu*, vol. 47 (230), no. 5–6 (330–331).
- Richard Lepsius (1863) *Standard Alphabet for Reducing Unwritten Languages*.
- Josip Lisac (2009) *Hrvatska dijalektologija 2: Čakavsko narječje*. Golden marketing – tehnička knjiga, Zagreb.
- Keļmakov (2004): В.К. Кельмаков, *Диалектная и историческая фонетика удмуртского языка*.
- Ludwig Erich Schmitt (1964) *Zeitschrift für Mundartforschung*. Wiesbaden: Franz Steiner.
- A.N. Tucker (1971) Orthographic systems and conventions in Sub-Saharan Africa. In Sebeok (ed.) *Current trends in linguistics, volume 7: Linguistics in Sub-Saharan Africa*. Mouton.

Figures

MINĐA **m'inža m'inže** *f ženski spolni organ*. ♦ J'ona je d'obru zn'ola sv'oju m'inžu skur'istitj duk je ml'ajša b'ila. V'e ju v'ise m'uški neti ne gled'iju. (U Prelogu nije zabilježena ta imenica, a u značenju ženskog spolnog organa upotrebljava se imenica c'uca.) (f)

MOČJENK **m'očjenk m'očjenka** *m jelo od kukuruznog brašna*. (U Prelogu je m'očnak. U Svetom Đurđu je došlo do prelaska *ń* u oblik s anticipiranim *j* u slijedu *jn*, a zanimljivo je da se *e* iz tvorbenog nastavka pomaklo između *j* i *n*.) (a)

Figure 1. Belović & Blažeka (2015: 48). The fused diacritics on č and ž.

ž – glas između ž i ž
 č – glas srednji između č i ć

Figure 2. Lisac (2009: 10). This reads, “ž [is] a sound between ž and ž” and “č [is] a sound intermediate between č and ć,” where ž and c have the phonetic values [dz] and [ts].

ž/z'	ž'
ž/ž'	ž
č/č'	ч

Figure 3. Keľmakov (2004: vol. 1, p.13 and vol. 2, p. 11). A Latin–Cyrillic correspondence table that presents palatalized Latin letters such as *ń*/*n'* and *ś*/*s'* in allographic pairs. When it comes to the crown letters, *ž* and *č* are equivalent to *ž'* and *č'*, demonstrating that the middle tine of the crown is identified as a palatalization mark, with the rest being a caron.

schwach vorgestülpt werden. R. MARTIN¹²⁹ schreibt zu diesem Komplex: „Es bestehen zwischen *s* und *š* eine lange Reihe Zwischenstufen, die nur an Ort und Stelle im Laufe der Darstellung geschildert werden können.“ Auch für den Übergang zum guttural-palatalen Bereich wurden einige Zwischenwerte gefunden: PROTZE¹³⁰ charakterisiert $\frac{\downarrow}{s}$ als stimmlosen alveolaren Zischlaut, zwischen *š* und *χ*, für GROSSE¹³¹ ist *χ'* ein stl. palataler, leicht zischender Reibelaut (Zwischenstufe zwischen *š* und *χ*).

Figure 4. Schmitt (1964) in L2/11-202. The textual illustration that was used for U+1AB3 in its Unicode proposal.

The series of *palatal* and *palato-dental* sounds will therefore be as follows:
k g n; χ γ; š ž; ṣ̌ ẓ̌; y; l.

Figure 5. Lepsius (1863: 73). The discrete, but kerned, diacritics in *ṣ̌* and *ẓ̌*. Note the high angle of the acute accent even in isolation. Lepsius set all phonetic notation in italic typeface.

Alveolar voiced					
fricative	<i>z</i>	<i>z</i>	<i>z</i>	<i>z</i>	<i>z</i>
[retroflex		<i>ẓ</i>	<i>ẓ</i>	<i>ẓ</i>	<i>ẓ</i>]
palatalized	<i>zy</i>	<i>ẓ</i>	<i>ẓ</i>	<i>ẓ</i>	
labialized	<i>ẓ</i>			<i>ẓ</i>	<i>ẓv, ẓw, ẓ</i>
[velarized		<i>ẓ, ɔ̣</i>	<i>ẓ</i>	<i>ẓ</i>	<i>ẓ</i>]

Figure 21. Z-swash tail (red highlight) in contrast with z-retroflex hook (blue highlight); from Tucker (1971), p. 648.

Figure 6. Tucker (1971: 648). Note the crown diacritic at right in this figure from [L2/03-190](#). The orthography is Meinhof's.