Unicode request: Dania (Danish dialectological) symbols

Kirk Miller, draft, 2020 jan 10
chart from: Ruben Schachtenhaufen <ruben@schwa.dk> (date: 2017 aug 15)

Needed characters (previously proposed for Teut(h)onista):
ɟ ꝅ ɭ ɟ   ⱱ ꭉ ʎ

Potential need, to be evaluated by the Unicode committees:
a' a̗ 0 ⌐ ɷ ʒ æ (italic ɑ-e ligature, distinct from æ and œ)

Jespersen (1890) and Kristensen (1924) have additional characters which are not included in the Dania computer font used for this document. For Jespersen, there are roman m n b d g w l. Kinstensen adds a turned ⱱ and roman v ð q. The roman letters stand out from the typeface, which is by default italic as an indication of phonetic transcription. Early versions of the Teuthonista proposal included *left*-inclined letters to handle most of these. Since they are not found in the Dania computer font that has been shared with me, they may be obsolete.

Most of these characters (but not the Krisensen additions) were included in early versions of the Teut(h)onista proposal to Unicode, but dropped without any explanation that I can find.

Konsonanter.

	Lukke		Midtaabning		Side-aab-ning	Snur-ren
	Næselyd	Mundlyd	tynd	bred		
Læber	m *m*	*p b* b	w *ƕ w*	— *ꝺ*		
Underlæbe	— *ɱ*		*f v*			
Tungespids . . . {	n *n*	*t d* d	*s z*	*þ ð*	l *l; t*	*ꭆ ꭆ*
	— *ṇ*	*ṭ —*	*ç ʒ̣*	— *r*	— *ḷ*	
Do. og fortunge	— *ꞑ*	*ṱ ꝱ*	*ṣ —*	— (*ð*)	— *ḻ*	
Fortunge	— *ɲ*	*ʄ ɟ*	*ʃ ʒ*	*c ɧ j*	— *ɟ*	
Midttunge	— (*ɲ*)	*ꝅ ɡ*		*χ ɣ*		
Bagtunge	— *ŋ*	*k g* g		*x q*		*ⱱ ꭉ*
Tungerod				*ᴚ r*		
Strube		*' —*		*h —*		

[I hver rubrik staar først den stemmeløse og sidst den stemte lyd.]

Figure 1. Jesperen p. 40, consonants.

A) Red arrows indicate letters that remain roman despite the use of italic typeface to indicate Dania transcription. Kinstensen adds roman v ð q.

B) Yellow arrows indicate novel letters. The voiced palatal fricative ʒ̢ might be handled in Unicode by U+293 ʒ. (They are graphically closer in the table than in the current computer font.) The closest characters in Unicode to the other consonants with a loop, ƫ ꝱ ꝅ ɭ, are not very similar. Kinstensen adds a turned ⱱ.

Vokaler.

Fortunge		Midttunge		Bagtunge	
Urundede	Rundede	Urundede	Rundede	Urundede	Rundede
i/i	y/y	?	$\ddot{u}, \dot{u}$		u/g
e/ε	$\emptyset/\varrho$	∂	$\dot{o}$		$o,0/\infty$
⇨ $\alpha/\mathit{æ}$	⇨ $\theta/\ddot{o}$	$\mathit{3}$	$\mathit{5}$	$\dot{a}$ α	$\mathring{a}/\mathit{o}$
$\ddot{a}$	$\mathit{5}$	$\mathit{v}, \mathit{\lambda}$		a	

Figure 1. Jesperen p. 40, vowels. It is not clear if these require Unicode support.

A) The front unrounded vowels make a distinction between fully italic 'æ' and a merely inclined variant, analogous to the a–ɑ distinction in IPA. The variant at left might be handled by 'œ', which does not occur. (Note that the tick at the top center of the œ-like letter shows that it is an italic 'æ', not 'œ'. Similarly, the mid back vowel ꜵ appears to be U+A735 ꜵ rather than U+A74F ꝏ.) Correspondence with users of Dania, however, suggests they feel it is important to keep æ distinct from 'œ'.

B) The front rounded vowel, ꝋ, might be handled with U+A74B ꝋ, if the inclination of the cross-bar can be expected to be handled by the design of the font.

C) The back vowel 0 appears to merely be the digit zero. The question is then whether a character defined as a digit can properly function as a phonetic symbol.

Diacritics

The 'nasal curl' a_c is semantically equivalent to an ogonek, and might be handled as such. However, graphically it resembles a subscript 'c', which I will be requesting in a separate Unicode request.

$a^{\prime}$ would seem to be a simple combination of the length sign $a^{\cdot}$ with stød $a^{\prime}$, and if so would not require Unicode support.

⌐ is included in the Dania font but not defined in the tables I've seen. A modifier letter open corner ◌˹ is however found in other phonetic traditions, and I will be requesting it separately as well.

The Dania font includes a combining descending loop diacritic, but since it intersects with the base letter and all combinations would therefor presumably receive their own code points, I don't know if a separate Unicode point for the diacritic itself would be called for.

Table. Basic Dania Latin characters not equated with Unicode in Schachtenhaufen 2017

Following are those characters in Schachtenhaufen's chart 'signographie.de & Ømålsordbogen: Latin Modern Dania (prøvefont version 8)' not there equated with Unicode, ignoring Greek letters and precombined letter-diacritics. Those which are not easily equated with Unicode are highlit.

No.	Character	Name	Unicode equivalent
13	æ	LATIN SMALL LETTER DANIA AE	= 'æ'? [1]
27	ɖ	LATIN SMALL LETTER D WITH LOOP	
46	ɡ	LATIN SMALL LETTER SCRIPT G WITH DESCENDING LOOP	AB36 (German dialectology)
47	ɋ	LATIN SMALL LETTER SCRIPT G WITH HOOKTAIL	24B or just ɋ [2]
58	ɩ	LATIN SMALL LETTER DESCENDING DOTLESS I	A781 [3] (+ dot for no. 59 ɩ̇)
65	ƙ	LATIN SMALL LETTER K WITH DESCENDING LOOP	
70	ɭ	LATIN SMALL LETTER L WITH DESCENDING LOOP	
72	ɯ	LATIN SMALL LETTER M ROTUNDA	28D [4]
76	ŋ	LATIN SMALL LETTER N WITH DESCENDING LOOP	AB3B, AB3C (German dial.)
85	θ	LATIN SMALL LETTER ASCENDING O	0? (~ theta without a bar) [5]
87	ȣ	LATIN SMALL LETTER LITTLE OU	223
94	ⱺ or ɵ	LATIN SMALL LETTER O WITH REVERSED STROKE	1D13 or A74B? [6]
104	ɹ	LATIN SMALL LETTER SCRIPT R WITH LEFT ONSET	
105	ɼ	LATIN SMALL LETTER R WITH LEFT SWASH	
106	ℛ	LATIN SMALL LETTER CAPITAL R WITH LEFT SWASH	
107	ɾ	LATIN SMALL LETTER SCRIPT R	AB4B (German dial.)
116	ƫ	LATIN SMALL LETTER T WITH DESCENDING LOOP	
124	ʌ	LATIN SMALL LETTER TURNED V	28C
134	ɥ	LATIN SMALL LETTER Y WITH STRAIGHT DESCENDER	265 [7]
136	ʒ (ȝ)	LATIN SMALL LETTER Z WITH DESCENDING LOOP	293? [8]
141	᾿	DANIA COMMAPUNKT	(simple combination) [9]
143	·	DISTINCTIO	B7 (assumed)

---

1. Is the *a-e* ligature actually in a 3-way contrast with ⟨æ⟩ and ⟨œ⟩? The Dania font only has a 2-way contrast.
2. U+24B (ɋ) is defined in Unicode as 'Latin small letter q with hook tail'. However, there would seem to be no graphic difference between the two, and a simple substitution should be workable. Given that the roman variant is simply 'q', however, it may be that italic *q* is all that is needed, though the current Dania font does distinguish italic *q* from ɋ.
3. U+A781 is 'turned l', but graphically should be identical to the stretched (dotless) i.
4. In the Dania font, the italic 'm rotunda' is simply a turned italic *w*, so presumably Unicode turned *w* should work. The shapes are parallel to ʌ vs *v* in italic typeface, and differ from Gentium only in the graphic design of the font.
5. A separate code point might not be needed if digit zero works as a substitute.
6. U+1D13, 'sideways o with stroke', is in the Uralic Phonetic Alphabet for reduced [ø], but it's inclined the wrong way.
7. There is also U+2AE ɥ and 2AF ɥ , but they have in-built swashes (ɥ ɥ) that may not be appropriate for straight *y*.
8. U+293 is defined as 'ezh with curl', and may or may not be a good fit graphically.
9. Shown here is interpunkt + combining apostrophe, 2E + 313. Cf. also 2E4C medieval comma.

145	ą (ą)	NASAL CURL	328 ogonek? subscript c?
160	‖	DOUBLE VERTICAL STRONG BAR	1D12 (Teutonista decision)
161	‖	DOUBLE VERTICAL THIN BAR	2016
162		DANIA LOOP DESCENDING	(no need for Unicode point?)

References

Otto Jespersen, *Dansk lydskrift*, in *Dania*, vol. 1 (1890-1892), pp. 33-79.
   http://jysk.au.dk/fileadmin/www.jysk.au.dk/publikationer/centrets_publikationer/dania.pdf
Marius Kristensen, *Vejledning til brugen af Danias lydskrift*, Copenhagen, 1924.
   http://jysk.au.dk/fileadmin/www.jysk.au.dk/publikationer/centrets_publikationer/vejledningdania.pdf