

On LATIN SMALL LETTER TURNED O OPEN-O and LATIN SMALL LETTER TURNED O OPEN-O WITH STROKE

Denis Moyogo Jacquerye <moyogo@gmail.com>

2022-06-25

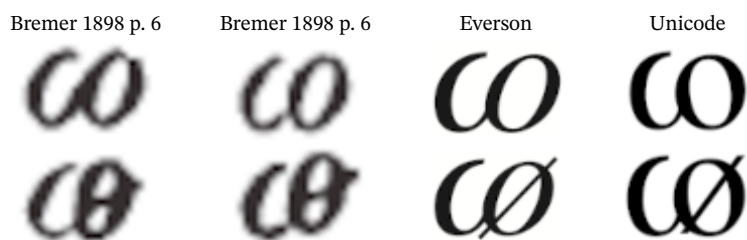
This document supersedes L2/22-097 and addresses points raised in L2/22-109.

The two characters U+AB43 LATIN SMALL LETTER TURNED O OPEN-O and U+AB44 LATIN SMALL LETTER TURNED O OPEN-O WITH STROKE were proposed for encoding in *Revised proposal to encode “Teuthonista” phonetic characters in the UCS, N4081, L2/11-202*.

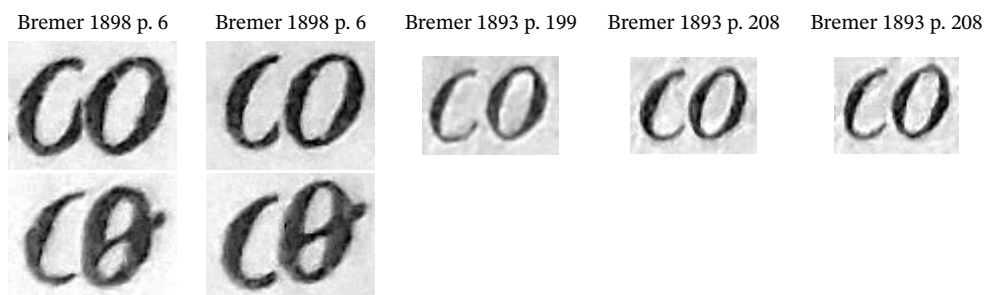
Their character names and representative glyphs (connected at the bottom) are problematic because they are based on approximate shapes from the following low quality high contrast sample from Bremer 1898, p. 6:

Mein bisheriges ð e € æ, ö o ɔ œ, ø θ ʃ ʈ
 ersetze ich jetzt durch e € æ a, o ɔ œ a, θ ʃ ʈ a
 und ebenso die Längen é ē ē ē, ó ô õ ö, ʈ ̈́ ̈́ ̈́
 durch ē ē ē ē, ô õ œ a, ̈́ ̈́ ̈́ ̈́

Here are the character glyphs from that Bremer 1898 p. 6 sample, representative italics from Everson’s L2/22-109 and Unicode’s representative glyphs in roman:



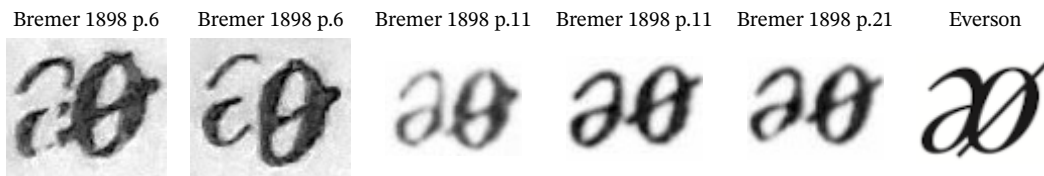
One may imagine ligatures of turned o+ɔ and turned ø+ɔ. That is not, however, what was set in type and published. Here are the characters from better quality samples in the same Bremer 1898 p. 6, in Bremer 1893 p. 199 and 208:



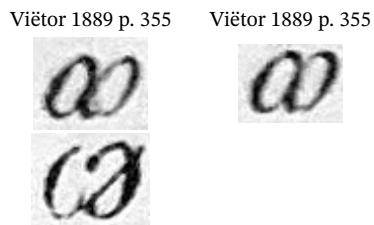
The actual glyphs in Bremer 1898 and Bremer 1893 have left parts that have been vertically cut from metal characters (most likely *o*). All are disconnected from the right parts, even one very lightly as the left part has been less generously cut. Not a single one looks like the current representative glyphs which have both parts connected at the bottom.

If one imagines the top was disconnected on purpose, one can imagine the bottom was disconnected on purpose as well. If one imagines the bottom is meant to connect, one can imagine the top is meant to connect. These shapes alone do not indicate that the typesetters formed these characters as ligatures turned *o*+*o* and turned *ø*+*o*.

Because of the inconsistencies, one should infer that Bremer 1898 had printing issues, especially when considering the glyphs of æ U+AB41 LATIN SMALL LETTER TURNED OE WITH STROKE in the sample from Bremer 1898, p. 6, along with selected glyphs from p. 11 and p. 21 in comparison with the representative italic glyph from Everson's L2/22-109:



In Bremer 1893, Otto Bremer describes the two symbols U+AB43 and U+AB44 as ligatures composed of *a* (italic script *a*), with *o* for the first and with *ø* for the second. In Viëtor 1889, the *a*+*o* shape is as described by Bremer, apparently composed of *a* and the cut *o* attached to it, and an approximate *a*+*ø* shape is used, where both right and left parts have been cut off where they should connect according to Bremer's description:



Bremer's alphabet did not use œ U+A74F LATIN SMALL LETTER OO. It used slanted double-storey *a* and script *a* contrastively. In later works using Bremer's alphabet, the symbol æ (æ U+1D14 LATIN SMALL LETTER TURNED OE) is replaced with æ (æ U+A735 LATIN SMALL LETTER AO) in Brandt 1913 and Schübel 1913 and æ (U+AB42 LATIN SMALL LETTER TURNED OE WITH HORIZONTAL STROKE) is replaced with a ligature slanted double-storey *a*+*ø* (not encoded) in Brandt 1913. These later forms match Bremer's 1898 description of the two symbols as ligatures of slanted double-storey *a* with *o* and with *ø* (the one with *ø* being implied). Bremer seems to have used turned œ as a practical solution and also applied the form to the capital of æ (Æ). Interestingly, Bremer used *A* and *æ* as the capital forms of *a* and *a* respectively.

Transcription systems using script *ao*

Noreen 1877, Landsmålsalfabetet in Lundell 1879, the Dania transcription in Jespersen 1890 and Noreen 1903 use a script *a+o* ligature. It is most likely borrowed from the italic form of the *æ* ligature (also used in Adolf Noreen's Old Norse work). Landsmålsalfabetet directly borrowed the symbol from Noreen 1877 and Dania borrowed it from Noreen 1877 or Landsmålsalfabetet. Eriksson 1961 p. 121 explicitly describes Landsmålsalfabetet's symbol as composed of *a* and *o*.

Because these systems use slanted double-storey *a* and script *a* contrastively, or use slanted double-storey *æ* and script *æ* (not encoded) contrastively, the glyph of *æ* U+A735 LATIN SMALL LETTER AO may not be adequate in italic fonts and a yet to be encoded character could be more appropriate, as fonts may use slanted double-storey *a* as a base for ligature characters composed with U+0061 LATIN SMALL LETTER A.

It must be noted that the 1924 version of the Anthropos alphabet presented in Heepe 1928 by Wilhelm Schmidt, which had symbols encoded in Everson *et al.* L2/11-202, uses slanted double-storey *æ*, slanted double-storey *æ* and slanted double-storey *æ* (*æ* U+AB31 LATIN SMALL LETTER A REVERSED-SCHWA) while also using *œ* (*œ* U+0153 LATIN SMALL LIGATURE OE) and *œ* (*œ* U+AB40 LATIN SMALL LETTER INVERTED OE). Italic fonts that do not use a slanted double-storey *a* to compose those ligatures of *a* are most likely inadequate for the Anthropos alphabet.

Form of script *ao*

The characters *a* U+0061 and *ɑ* U+0251 are inherently confusable in italic fonts where they have the same or similar shape because of the encoding model. This potential confusability is inherited by characters derived from or composed with them.

In the systems using slanted double-storey *a* and script *a* contrastively, the form of *æ* (*æ* U+A735 LATIN SMALL LETTER AO) would not be confusable with the script *ao* ligature. For example (*a* U+0061, *ɑ* U+0251, *æ* U+A735, script *ao*):

aaææ

The Greek-alpha form of the italic *ɑ* (*ɑ* U+0251) is not appropriate for those systems, just as the script *a* (*a* U+0251) form is not appropriate for the Uralic Phonetic Alphabet. Italic fonts that use the Greek-alpha form for *ɑ* U+0251 could have distinctive glyphs using shapes similar to the following for (*a* U+0061, *ɑ* U+0251, *æ* U+A735, alpha-o):

ααωω

Similar forms can be used for the variants *æ* U+00E6, slanted double-storey *æ* and yet to be encoded script *æ* used contrastively in the previously mentioned transcription systems.

Representative glyph recommendations

The representative glyphs of U+AB43 and U+AB44 should be changed given that:

- the current representative glyphs, connected at the bottom, do not correspond to the glyphs, disconnected at the bottom, seen in Bremer's works;
- Bremer 1893 and Bremer 1898 had printing artifacts: U+AB43 and U+AB44 are composed of inconsistently cut or positionned parts.

U+AB43 and U+AB44 should have been ligatures composed of script *a* given that :

- Viëtor 1889 establishes the shape of U+AB43 as Bremer 1893 describes it (ligature composed of *a* and *o*);
- Bremer's alphabet differentiated slanted double-storey *a* from script *a*, and symbols composed with them by using double-storey *a* or *æ* for symbols composed with *a* and by using script forms for symbols composed with *a*.
- Bremer 1898's *æ* (*a* with stroke) was encoded as *æ* U+2C65 LATIN SMALL LETTER A WITH STROKE composed of a U+0061 LATIN SMALL LETTER A because of its shape;
- Bremer 1898's *α* (barred *a*) was encoded as *α* U+AB30 LATIN SMALL LETTER BARRED ALPHA, composed of *α* U+0251 LATIN SMALL LETTER ALPHA because of its shape.

The ligature script *ao* symbol could be encoded as an unambiguous character distinct from U+AB43 because:

- phonetic transcription systems (Noreen 1877, Landsmålsalfabetet, Dania) use a ligature script *ao* symbol;
- these systems use *a* and *æ* contrastively, like Bremer's;
- the slanted double-storey *æ* form of U+A735 LATIN SMALL LETTER AO is not appropriate for those systems;
- the disconnected parts form of U+AB43 is not appropriate for those systems.

This character is not being proposed for encoding in this document.

The representative glyphs of U+AB43 and U+AB44 should either be changed to correspond to disconnected parts as seen in Bremer 1893 and Bremer 1898, or to ligatures composed of script *a* as seen in Viëtor 1889.

Because of the error in the character names of U+AB43 and U+AB44, which cannot be changed, the disconnected form (with disconnected top and bottom) is recommended at this point. The following representative glyphs are proposed:

CO CØ

For reference, the following glyphs could be used for italic.

Proposed update to UTN #27

The *Unicode Technical Note #27: Known Anomalies in Unicode Character Names* should be updated with the following:

U+AB43 LATIN SMALL LETTER TURNED O OPEN-O

- Intended for a ligature script a-o but rendered as a disconnected cut o and o.

U+AB44 LATIN SMALL LETTER TURNED O OPEN-O WITH STROKE

- Intended for a ligature script a-ø but rendered as a disconnected cut o and ø.

Bremer's alphabet samples

Vokalzeichen: a, ω, ɔ, o, ǫ, u, ŭ; œ, ɶ, ø, ǿ, y, ȳ normale zungenbreite
(ô, ô, o, ǫ, û ŭ;) æ, ɛ, e, ẽ, i, ĭ verbreiterte hinterzunge.

Figure 1. Bremer's symbols for vowels as presented by Wilhelm Viëtor in *Phonetische Studien*, vol. 2, 1889, p. 355. showing co U+AB43 as, script ω.

wenn wir statt *au* etwa *acooou* schreiben wollten, so hätten wir nur 5 bestimmte Punkte aus der Vokalskala gekenn-

Figure 2. Bremer 1893, p. 199 showing co U+AB43 with disconnected parts.

Das· îx̂· | ər.v:áxt· || cou̇s. mæ̇i.nər. šḋi.lən. hỳ·tə· |
Dən. bérk· hi.ncou̇f· | mit. frí·šər· zé':lə. gìṅk. ||

Figure 3. Bremer 1893, p. 208 showing co U+AB43 with disconnected parts.

Soweit meine Kenntniss der deutschen Mundarten reicht, ist es notwendig, dass man vom engen *e*, *o* und *ö* hin bis zum reinen *a* je drei Vokalstufen unterscheide.

Mein bisheriges. *ë e € æ, ö o ɔ oo, ø θ ʃ œ*
 ersetze ich jetzt durch *e € æ a, o ɔ œ a, θ ʃ œ a*
 und ebenso die Längen *é ē ē ā, ó ō ō ō, ê ē ē ē*
 durch *ē ē ā ā, ō ō ā ā, ē ē ē ē*

Figure 4. Bremer 1898 (1969 reprint), p. 6 showing co U+AB43 with disconnected parts.

bet. *æ, œ* und *œ* sind die Ligaturen von *a* (*a*) und *e*, *o* und *ø*, wie sie die Mittellaute zwischen *a* und *e*, *o* und *ø* ausdrücken sollen. *a, a* und *æ* erkennt man sofort als eine Abart des *a*: diese Buchstaben bezeichnen Vokale, die man kaum noch überweite *e*, *o* und *ø* nennen darf, es sind vielmehr die Buchstaben für die *a*-Laute, die ein wenig *e*-, *o*- und *ø*-artig klingen, im Gegensatz zu dem reinen *a*.

Figure 5. Bremer 1898, p. 7 describing *æ*, *œ* and *œ* as ligatures of *a* (*a*) with *e*, *o* and *ø*.

in unbetonter Silbe und ist je nach der Umgebung bald etwas heller, bald etwas dunkler: *drivœ* treiben.

§ 43. *œ* kommt nur im Diphthong *œu* vor (siehe § 51).

§ 44. *ā* ist weites überlanges *a*, das dem bühnendeutschen und italienischen langen *a* entspricht: *hāvsth* Herbst (vgl. auch § 48 und 92, 1a und 2a).

Figure 6. Brandt 1913, p. 37, one of the few of works showing *œ* U+AB42 rendered as slanted double-storey *œ* (glyph variant or character yet to be encoded).

§ 45. *ā* ist ein weites langes *a*, das dem bühnendeutschen langen *a* nicht völlig gleich ist, sondern etwas, wenn auch nur wenig nach *œ* hin davon abweicht: *gāsth* Geest.

§ 46. *ā* ist der weite lange Vokal, der in der Mundart von M dem *ā* von N und S entspricht. Der Klangcharakter gleicht dem Vokal im engl. *to fall*: *gāsth* Geest M.





Figure 7. Brandt 1913, p. 38, one of the few of works showing *œ* U+1D14 rendered as slanted double-storey *œ* U+A735.

„Gregorius“ (Ausgabe PAUL) beweist, und von denen die weder stamm- noch endungsbetonten noch heute in der Schriftsprache auf der Mittelsilbe betont sind. So heißt es: *ædſ* Adam, *Æixid* Ägidius, *Mæri* Maria, *Mæriaſid* Marienlied in der katholischen Kirche, *Mæds* (< *Matthias*).

Figure 8. Schübel 1913, p. 22, one of the few of works showing æ U+1D14 rendered as slanted double-storey æ U+A735.

Other samples

Vokaler.

Fortunge		Midttunge		Bagtunge	
Urundede	Rundede	Urundede	Rundede	Urundede	Rundede
<i>i/i</i>	<i>y/ɥ</i>	<i>ɨ</i>	<i>ü, ù</i>		<i>u/ɤ</i>
<i>e/ɛ</i>	<i>ø/ɘ</i>	<i>ə</i>	<i>ó</i>		<i>o, 0/ɔ</i> 
<i>æ/æ</i> 	<i>ø/ö</i>	<i>ɜ</i>	<i>ǝ</i>	<i>á</i> 	<i>å/ɔ</i>
<i>ä</i>	<i>ǟ</i>	<i>ɐ, ɹ</i>		<i>a</i> 	

[Over skraastregen staar „tynde“, under den „brede“ vokaler.]

Figure 9. Jespersen 1890, p. 40 showing slanted double-storey *a* and script *a* used contrastively, slanted double-storey *æ* and script *æ* used contrastively, and ligature script *a+o* in the Dania phonetic alphabet.

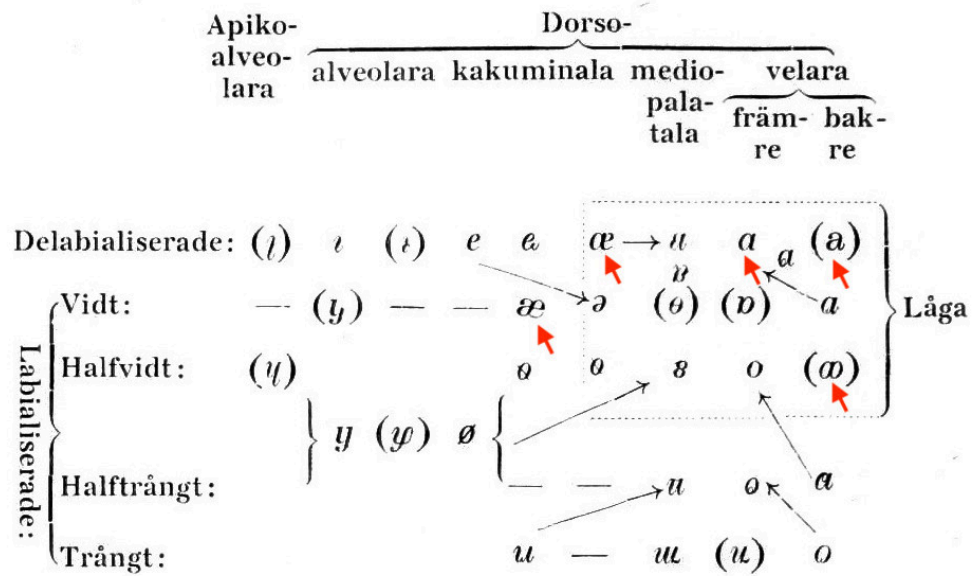


Figure 10. Noreen 1903, p. 538, showing slanted double-storey *a* and script *a* used contrastively, slanted double-storey *æ* and script *æ* used contrastively, and ligature script *a+o* in Landsmålsalfabetet.

a = a, a, <i>o</i> , (a)	t = t, t, f
b = b, (b)	u = u, u, u
d = d, d, d, d	v = v
e = e, e, e, (i, æ)	w = w
f = f	y = y, y, y, (y)
g = g, g, g	å = a, (o, o)
h = h, (m, n, g, h, o)	ä = æ, a, (e)
i = i, i, i, (j)	ö = ø, ø, (x, ø)
j = j, j, j, j, (i)	rt = rt, rt, t, t
k = k, k	rd = rd, rd, d, d
l = l, l, l, l, l, J, (j)	rl = rl, l
m = m, (m)	rn = rn, rn, n, n
n = n, n, (n, y)	rs = rs, rs, s, s
g = g, (g)	lt = lt, lt, lt, t, t
o = o, ø, ø, (a, u)	ld = ld, ld, d, d
o = o, (u)	ln = ln, ln, n, n
p = p	ls = ls, ls, s, s
r = r, r, r, n, (r, r)	tj = tj, tj, js, j, s
s = s, z, l, (s, j)	nj = nj, y.
ſ = f, f, s, s, (s)	

Figure 11. Eriksson 1961, p. 133 showing ligature script *a+o* and forms of U+00E6: slanted double-storey *æ* and script *æ* used contrastively, in Landsmålsalfabetet.

DIE ANTHROPOS-LAUTSCHRIFT¹⁾

Einteilung der Vokale.

[illegible]

Figure 12. Schmidt 1928 (in Heepe 1928), p. 6 showing italic *a*, slanted double-storey æ U+00E6, slanted double-storey æ U+A735, slanted double-storey æ U+AB31, as well as œ U+0153 LATIN SMALL LETTER OE and œ U+AB40 in the 1924 version of the Anthropos alphabet.

References

Bremer's alphabet

- Ernst Brandt, *Die nordfriesische Sprache der Goesharden*, Halle a.d.S., Waisenhauses, 1913. <https://babel.hathitrust.org/cgi/pt?id=uc1.a0011731866>
- Otto Bremer, *Deutsche Phonetik*, Leipzig, Breitkopf und Härtel, coll. « Sammlung kurzer Grammatiken deutscher Mundarten » (no. 1), 1893. <https://archive.org/details/deutschephoneti00bremgoog> or <https://katalog.ub.uni-leipzig.de/Record/0-1657859460/Description>
- Otto Bremer, *Beiträge zur Geographie der Deutschen Mundarten in form einer Kritik von Wenkers Sprachatlas des deutschen Reichs* 1895. https://archive.org/details/bub_gb_A3sKAAAAIAAJ
- Otto Bremer, *Zur Lautschrift*, Leipzig, Breitkopf und Härtel, coll. « Sammlung kurzer Grammatiken deutscher Mundarten » (no. 1 Anhang), 1898. https://archive.org/details/bub_gb_pTkTAAAAQAAJ or <https://katalog.ub.uni-leipzig.de/Record/0-1657873358/Description>
- Manne Eriksson, *Svensk ljudskrift 1878–1960 : En översikt över det svenska landsmålsalfabetets utveckling och användning huvudsakligen i tidskriften Svenska Landsmål*, Stockholm, P. A. Norstedt & Söner, 1961.
- Michael Everson, Alois Dicklberger, Karl Pentzlin and Eveline Wandl-Vogt, *Revised proposal to encode “Teuthonista” phonetic characters in the UCS*, N4081, L2/11-202), 2011-06-02. <https://www.unicode.org/L2/L2011/11202-n4081-teuthonista.pdf>
- Michael Everson. *Exploratory proposal to encode Germanicist, Nordicist, and other phonetic characters in the UCS* N3555, L2/08-428, 2008-11-27. <https://www.unicode.org/L2/L2008/>

08428-n3555-loudstaves.pdf

- Michael Everson, Eveline Wandl-Vogt, Alois Dicklberger. *Preliminary proposal to encode “Teuthonista” phonetic characters in the UCS* N3907, L2/10-346, 2010-09-23. <https://www.unicode.org/L2/L2010/10346-n3907-teuthonista.pdf>
- Michael Everson, Eveline Wandl-Vogt, Alois Dicklberger. *Proposal to encode “Teuthonista” phonetic characters in the UCS*. N4031, L2/11-137, 2011-05-09. <https://www.unicode.org/L2/L2011/11137-n4031-teuthonista.pdf>
- Otto Kürsten and Otto Bremer, *Lautlehre der Mundart von Buttstedt bei Weimar*, Leipzig: Breitkopf & Härtel, coll. “Sammlung kurzer Grammatiken deutscher Mundarten”, no. 9, 1910. <https://katalog.ub.uni-leipzig.de/Record/0-1657849023/Description>
- Ferdinand Herdemann, *Versuch einer Lautlehre der westmünderländischen Mundart*, Leipzig, 1921. <https://babel.hathitrust.org/cgi/pt?id=uc1.b5218900>
- Gesinus Kloeke, *Der Vokalismus der Mundart von Finkenwärder bei Hamburg*, Hamburg, 1913. <https://babel.hathitrust.org/cgi/pt?id=uc1.d0005981709>
- Gesinus Kloeke, *Der Vokalismus der Mundart von Finkenwärder bei Hamburg*, Hamburg, 1914. <https://babel.hathitrust.org/cgi/pt?id=uc1.b2631527>
- Hugo Larsson, *Lautstand der Mundart der gemeinde Altengamme (in den Vierlanden bei Hamburg)*, 1917. <https://babel.hathitrust.org/cgi/pt?id=uc1.b2643067>
- Georg Schübel, *Versuch einer Charakteristik und Phonetik der Bamberger Mundart von Stadtsteinach*, Halle a.d.S., Waisenhaus, 1911. <https://babel.hathitrust.org/cgi/pt?id=uc1.b2640575>
- *Unicode Technical Note #27*. <http://www.unicode.org/notes/tn27/>
- Wilhelm Viëtor, “Eine sammlung von grammatiken deutscher mundarten”, *Phonetische Studien*, vol. 2. 1889, p. 353-356. <https://archive.org/details/phonetischestudi01marb>

Other phonetic alphabets

- Otto Jespersen, “Danias lydskrift”, *Dania: Tidsskrift for Folkemål og Folkeminder*, 1890. http://jysk.au.dk/fileadmin/www.jysk.au.dk/publikationer/centrets_publikationer/dania.pdf
- Johan August Lundell, “Det svenska landsmålsalfabetet”, *Svenska Landsmål*, no. 1, 1878, p. 11–158. https://www.isof.se/download/18.317326fe179568897271b050/1621945762141/Svenska%20landsm%C3%A5l%20och%20Svenskt%20folkkliv_1878_h1.pdf
- Adolf Noreen, *Vårt språk 1*. Lund: Gleerup, 1903. <https://archive.org/details/vrtsprknysvenskg01nore>
- Adolf Noreen, *Fryksdalsmålets ljudlära*, Upsala: E. Edquists boktryckeri, 1877. <https://babel.hathitrust.org/cgi/pt?id=wu.89099895609>
- Wilhelm Schmidt, “Die Anthropos-Lautschrift”, in Martin Heepe, *Lautzeichen und ihre Anwendung in verschiedenen Sprachgebieten*, Berlin: Reichsdruckerei, 1928, p. 6-10.