Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

1. Introduction. A set of characters used by specialists in medieval European philology, palaeography, and linguistics, as well as by other specialists such as Iranianists, is absent from the Universal Character Set. One of the persistent difficulties in encoding punctuation in the UCS is the fact that, in many traditions, dots are arranged in various orientations and configurations. UTC feedback on N3178 (a proposal to encode the Avestan script, including punctuation) requested that a review of the punctuation there be undertaken and that a comprehensive set of recommendations be made.

2. Review of existing "generic" punctuation. The core definition of "generic" punctuation is based on the typical shapes and sizes of punctuation as used in the Latin, Greek, and Cyrillic traditions; Georgian may be added to this list. When such punctuation is used in other scripts, such as the scripts of India or Southeast Asia, they are often modified to suit the ductus and letterforms of those scripts, but nevertheless the core definitions are based in the habits of European typography. An examination of the height of the modern : COLON U+003A, along with the size of its dots, is probably the best way to compare the existing set of punctuation characters and to determine how specialist needs can be met by filling out gaps or by defining characters within the context of the set of punctuation characters already encoded in the UCS.

2.1. Low or baseline punctuation. The characters , COMMA U+002C and . FULL STOP U+002E form the basis for the system.

2.2. Middle or x-height punctuation. The characters : COLON U+003A and ; SEMICOLON U+003B define the basic height for most inline generic punctuation dots. Conforming to this configuration are ⁏ REVERSED SEMICOLON U+204F, ⁖ GEORGIAN PARAGRAPH SEPARATOR U+10FB, ⁝ THREE DOT PUNCTUATION U+2056, ⁘ FOUR DOT PUNCTUATION U+2058, ⁙ FIVE DOT PUNCTUATION U+2059, ※ U+203B REFERENCE MARK and ⁜ DOTTED CROSS U+205C belong to this group. Missing are upward and downward pointing triangles.

2.3. High or tall punctuation. The characters ⁚ TWO DOT PUNCTUATION U+205A, ⁛ FOUR DOT MARK U+205B, ⁝ TRICOLON U+205D, and ⁞ VERTICAL FOUR DOTS U+205E all extend from baseline or just below it to caps height.

3. Medievalist additions. Although many editors substitute modern for medieval punctuation, a growing number of medievalists insist that modern punctuation cannot adequately represent the syntactical features of medieval texts. In any case, at present, the reproduction of the punctuation of most medieval manuscripts using currently defined UCS characters is impossible, because many marks (e.g. PUNCTUS ELEVATUS) are simply missing. More subtly, however, even those medieval marks that have modern reflexes may differ in usage as well as in shape. For example, the modern ? QUESTION MARK U+003F is descended from the ⸮ PUNCTUS INTERROGATIVUS, but unlike the modern QUESTION MARK, the PUNCTUS INTERROGATIVUS may be contrasted with the ꛷ PUNCTUS PERCONTATIVUS—the former marking questions that require an answer while the latter marks rhetorical questions. These medieval punctuation marks are proposed below. Note that where a number of characters are described as "variants of POSITURA", this

1

does not mean that they are freely-exchangeable glyph variants. This collection of punctuation marks is a superset of the letters found in the medieval corpus; no single manuscript contains all of them.

3.1. Low or baseline Medievalist punctuation. Some of the punctuation used by medievalists belongs to the "baseline punctuation" group.

, POSITURA is used to mark the end of a section of text. It is also used as the low SIMPLEX DUCTUS. (Figures 32, 33, 38)

; LOW PUNCTUS VERSUS is a variant form of PUNCTUS VERSUS. (Figure 21)

3.2. Middle or x-height Medievalist punctuation. Most of the punctuation used by medievalists belongs to the "x-height punctuation" group. A number of these are described as "variants of POSITURA". Parkes (p. 36) elaborates on this: "In western manuscripts the *positurae* fulfilled the need for more accurate indication of the nature of the pauses required to elucidate the sense of a text when it was to be intoned or sung in the liturgy.… The melodic formulae inevitably fell at the ends of sense units, thus the signs could come to be used as punctuation to mark pauses."

ˇ COMMA ELEVATA is used to indicate a brief medial pause. (Figures 29, 30)

ꝫ MEDIEVAL COMMA is used to mark a pause at the end of the rhetorical unit called a *comma*. (Figures 9, 22)

ꞏ PUNCTUS ELEVATUS is used in many medieval texts to signal a wide variety of sentence-medial pauses. (Figures 2, 6, 21, 22, 29, 30)

ꞏ PUNCTUS FLEXUS is used to mark sentence-medial pauses, but especially in liturgical texts where the pitch of the voice drops. (Figures 21, 22)

; PUNCTUS VERSUS is used to mark the end of a sentence or period, or the end of a verse of a psalm. (Figure 22)

:ˇ COLON WITH COMMA ELEVATA POSITURA, an alternative form of POSITURA, was "employed by insular scribes at the end of a paragraph in a series of paragraphs or texts (e.g. annals) to imply that some continuation was to be expected to complete this series". (Figures 26, 38)

:ꞌ COLON WITH RAISED COMMA POSITURA is a variant form of POSITURA. *Unfortunately, we are unable to find the evidence we had for this character when it was put into our font in 2005. Without a sample we may not be able to insist that the character be accepted, but we may point out that the character does fit into the pattern of other characters with a COMMA or a POSITURA element:* , ، ; ؛ :ˇ :ꝫ ꞌ ꝫ

:ꝫ COLON WITH POSITURA is a variant form of POSITURA. (Figures 22, 38)

∵ TWO DOTS OVER ONE DOT PUNCTUATION. (Figures 4, 22)

;ꞌ TWO DOTS OVER COMMA POSITURA is a variant form of POSITURA. (Figures 34, 38)

∴ ONE DOT OVER TWO DOTS PUNCTUATION can indicates a sentence break. (Figures 19, 22)

.ꞌ. THREE DOTS WITH COMMA POSITURA is a variant form of POSITURA. (Figure 38)

∷ SQUARED FOUR DOT PUNCTUATION. (Figures 7, 22)

⁙ FIVE DOT MARK is used to mark the end of a chapter or section in Old Slavonic. (Figure 18)

ꞌ MIDDLE COMMA is used along with middle dot for a certain kind of positura: ꞏꞌ . This is one of a set of positurae which can be composed with existing characters: .ꞌ and .ꞌ. and .ꝫ are examples. Note that :ˇ and :ꞌ and :ꝫ are not composable, because of the defined height of the COMMA ELEVATA, the MIDDLE COMMA, and the POSITURA: :ˇ and :ꞌ and :ꝫ (or even :ꝫ) are incorrect. The MIDDLE COMMA is also used as an abbreviation mark; for instance, when it follows long s, the reading is *sed* 'but'. (Figures 41, 42)

The collection of *positurae* here is a superset of the *positurae* found in the medieval corpus, comprising the most distinctive examples; no single manuscript contains all of them. At the same time, there is no single positura which does generic duty for this punctuation mark. The set here meets the needs of the greatest number of researchers.

2

3.3. High or tall Medievalist punctuation. Some of the punctuation used by medievalists belongs to the "tall punctuation" group.

⁊	SIMPLEX DUCTUS was (according to Parkes 1993, p. 307) originally a critical sign placed within a verse to separate matters accidentally run together; later it became a punctuation symbol. (Figure 32)

˙	DISTINCTIO is used (according to Parkes 1993, p. 13) to indicate "a final pause, after a *periodus*, or where the *sententia* is completed". The DISTINCTIO forms the third of a set of punctuation marks of increasing finality: the . FULL STOP, the · MIDDLE DOT, and the ˙ DISTINCTIO: . · ˙ (Figure 25)

⁒	PUNCTUS EXCLAMATIVUS (or *punctus admirativus*) is used in late medieval texts to mark the end of an exclamation. Iacopo Alpoleio da Urbisaglia claims to have invented it. (Figure 36, 39, 40)

⸮	PUNCTUS INTERROGATIVUS is used to mark the end of a question that requires an answer. (Figures 22, 25, 39, 40)

⸵	PUNCTUS PERCONTATIVUS is used in late texts to mark the end of a question that does not require an answer—a rhetorical question. (Figure 27)

∴	SIGNE DE RENVOI is used to associate matter in the text with material added in the margin, and especially a passage omitted from the text by the original copyist. It was later used as a quire mark by printers. (Figure 37)

/	VIRGULA INTERPOLATIVA is "used to mark the briefest pause or hesitation in a text". Petrarch used it where an interpolated statement had a different relationship to the general run of text. (Parkes 1993, p. 307; Figure 35)

Ⰳ	PARAGRAPHUS is used to mark the beginning of a section of text. (Figures 3, 5, 25)

◦	INTERPUNCTUM CIRCLE is used in transcriptions of epigraphy between letters or syllables or between words. (Figure 12b)

3.4. Editorial paired punctuation. The Nordic editorial tradition uses quilled bars to mark ⊦deletions⊣ in the manuscript and (in reverse order) to mark ⊣dittographies⊦ (repetitions). The venerable *Corpus Inscriptionum Latinarum* of Berlin makes use of a large number of paired punctuation, most of which are encoded in the General Punctuation or one of the Mathematics blocks. Two pairs are not yet encoded: the SIDEWAYS U BRACKETS are used to indicate Claudian letters in a transcription, to notate inverted letters, or to notate the insertion of an image into the line of text (such as ⊂crux⊃ to indicate a cross). The ⦅double parentheses⦆ are used by some other Latinists for similar purposes. They differ from ((doubled parentheses)) quite obviously, and are not joined at the top like the WHITE PARENTHESES U+2985-U+2986 are. The FULLWIDTH WHITE PARENTHESES U+FF5F-U+FF60 used in Asia are unusual in having either a double or a joined presentation, but that glyph variation is not used by Latinists.

⊦	LEFT VERTICAL BAR WITH QUILL (Figures 1, 8)

⊣	RIGHT VERTICAL BAR WITH QUILL

⌐	TOP LEFT HALF BRACKET (Figures 11, 12b)

¬	TOP RIGHT HALF BRACKET

∟	BOTTOM LEFT HALF BRACKET (Figure 10, 11, 43)

⌟	BOTTOM RIGHT HALF BRACKET

⊂	LEFT SIDEWAYS U BRACKET (Figures 12a, 12b)

⊃	RIGHT SIDEWAYS U BRACKET

⦅	LEFT DOUBLE PARENTHESIS

⦆	RIGHT DOUBLE PARENTHESIS

3.5. Modern grammatical punctuation. German dictionaries employ a variety of methods for demonstrating some morphophonemic processes. The ⸚ HYPHEN WITH DIAERESIS is used to indicate the umlaut of a stem vowel alongside a plural ending (so *Buch, ⸚er* means *Buch, Bücher*). The ~ TILDE is used in many dictionaries to avoid repetition of a unit (so *heiß, ~es* means *heiß, heißes*). Because German orthography writes nouns with an initial capital letter, and adjectives with a small letter, the retention or

change of case can be indicated with the dotted tildes (so *Buchmacher*, *≁markt*, *≁mäßig* means *Buchmacher*, *Buchmachermarkt*, *buchmachermäßig*). Some dictionaries use the ringed tilde to show the change of case (so *Buchmacher*, *~markt, ≗mäßig* means *Buchmacher, Buchmachermarkt, buchmachermäßig*). These punctuation characters are units, not equivalent to base characters plus diacritics.

⸚   HYPHEN WITH DIAERESIS (Figure 23)
≗   TILDE WITH RING (Figure 28)
⸛   TILDE WITH DOT ABOVE (Figure 24)
⸞   TILDE WITH DOT BELOW (Figure 24)

3.6. Comparison with mathematical symbols and punctuation. Some of the characters listed above are similar to characters used in mathematics. Now, it is the case that some characters with primarily mathematical uses are used by medievalists. Examples are:

⟦   27E6   MATHEMATICAL LEFT WHITE SQUARE BRACKET (Figures 11, 12a, 12b)
⟧   27E7   MATHEMATICAL RIGHT WHITE SQUARE BRACKET
⟨   27E8   MATHEMATICAL LEFT ANGLE BRACKET (Figure 10)
⟩   27E9   MATHEMATICAL RIGHT ANGLE BRACKET
⟪   27EA   MATHEMATICAL LEFT DOUBLE ANGLE BRACKET (Figure 12a)
⟫   27EB   MATHEMATICAL RIGHT DOUBLE ANGLE BRACKET

There is no reason that Medievalists should not use these paired punctuation characters, since the Medievalist use is also that of paired punctuation. But not all similar characters should be unified as these are. The following characters with mathematical uses are similar to characters proposed, but we do not propose a unification with them for several reasons.

∴   U+2234 THEREFORE should not be unified with ONE DOT OVER TWO DOTS PUNCTUATION because it is unrelated to the medievalist punctuation mark. Johann H. Rahn's *Teutsche Algebra* (Zürich 1659) uses both ∴ and ⸪ for THEREFORE. In Cajori (§668) the glyphs used for both of these is larger than the glyph of his COLON. Mathematicians have confirmed that this character should be caps-height. Since medievalist and generic punctuation is based on the colon, it is best to leave these mathematical characters to the mathemeticians, and to encode ONE DOT OVER TWO DOTS PUNCTUATION as supplemental punctuation, thus: : ⸫ ∴ vs ∴ ⸫ . The Sm (Symbol, Math) property applies to THEREFORE; the Po (Punctuation, Other) property applies to ONE DOT OVER TWO DOTS PUNCTUATION.

∵   U+2235 BECAUSE should not be unified with TWO DOTS OVER ONE DOT PUNCTUATION for the same reasons given for THEREFORE. It is definitely not a medieval character, and indeed Cajori (§669) says that it is not attested with the meaning BECAUSE before the 19th century, in the *Gentleman's Mathematical Companion* (1805). The Sm (Symbol, Math) property applies to BECAUSE; the Po (Punctuation, Other) property applies to TWO DOTS OVER ONE DOT PUNCTUATION.

∷   U+2237 PROPORTION should not be unified with SQUARED FOUR DOT PUNCTUATION because it was also devised by Gottfried Wilhelm Leibniz in 1696. In any particular font, the glyph for PROPORTION must be harmonized with the glyph for RATIO, while SQUARED FOUR DOT PUNCTUATION must harmonize with COLON. Further, as can be seen in Figures 7 (1470) and 22 (7th century), SQUARED FOUR DOT PUNCTUATION long predates PROPORTION and their origins are separate. (See Cajori §540.) The Sm (Symbol, Math) property applies to PROPORTION; the Po (Punctuation, Other) property applies to SQUARED FOUR DOT PUNCTUATION.

~   U+223C TILDE OPERATOR is not unified with TILDE. The glyph for TILDE OPERATOR ~ may differ from that of TILDE ~, though it applies to (at least) dozens of other mathematical characters, such as ∻ U+223B HOMOTHETIC, ≃ U+2243 ASYMPTOTICALLY EQUAL TO, ≈ U+2248 ALMOST EQUAL TO and ≋ U+224B ALMOST EQUAL TO. The proposed characters ≗ TILDE WITH RING and ⸞ TILDE WITH

DOT BELOW should not be encoded as mathematical characters, but as punctuation characters. The Sm (Symbol, Math) property applies to TILDE OPERATOR; the Sm (Symbol, Math) property also applies to TILDE, and to TILDE WITH RING and to TILDE WITH DOT BELOW. The mirroring bidi property applies to TILDE OPERATOR; it does not apply to TILDE, or to TILDE WITH RING or to TILDE WITH DOT BELOW.

≃ U+2A6A TILDE OPERATOR WITH DOT ABOVE should not be unified with ⩪ TILDE WITH DOT ABOVE, for the reasons given for TILDE OPERATOR. The Sm (Symbol, Math) property applies to TILDE OPERATOR WITH DOT ABOVE; the Sm (Symbol, Math) property also applies to TILDE WITH DOT ABOVE. The mirroring bidi property applies to TILDE OPERATOR WITH DOT ABOVE; it does not apply to TILDE WITH DOT ABOVE.

⊂ U+2282 SUBSET OF should not be unified with ⊂ LEFT SIDEWAYS U BRACKET. The height of the glyph for SUBSET OF may vary, but the height of the punctuation character does not; it is always low, so ⊂crux⊃ not *⊂crux⊃. The Sm (Symbol, Math) property applies to SUBSET OF; the Ps (Symbol, Math) property also applies to TILDE WITH DOT ABOVE. The mirroring bidi property applies to SUBSET OF; it also applies to TILDE WITH DOT ABOVE.

⊃ U+2283 SUPERSET OF should not be unified with ⊃ RIGHT SIDEWAYS U BRACKET, for the reasons given above.

The most difficult characters to evaluate are the HALF SQUARE BRACKETS. It is possible to consider unifiying them with either the QUINE CORNERS or the FLOORS and CEILINGS, used in mathematics. The problem here is that in mathematics, the specific shape of the QUINE CORNERS and the FLOORS and CEILINGS must be maintained, so that there is no confusion between them. The scholarly HALF BRACKETS have been found with some glyph variation, sometimes looking like QUINE CORNERS and sometimes like the FLOORS and CEILINGS; this is probably due to available fonts more than anything else—the "real" character is a half a bracket, or a bracket cut in half.

The ideally formed QUINE CORNER (U+231C-U+231F) has equal-length sides and sit at cap height and on the baseline: ⌞ ⌜ X ⌝ ⌟. Stroke width is uniform (not variable as for SQUARE BRACKETS and HALF SQUARE BRACKETS). See Figure 44.

The ideally-formed U+230A ⌊ LEFT FLOOR is shaped like tall L, or like the [with one of its horizontals filed off.) and its horizontal and vertical reflections appropriately translated about, with floor below the baseline and ceiling at about cap height: ⌊[X]⌋. Stroke width is uniform (not variable as for SQUARE BRACKETS and HALF SQUARE BRACKETS). The horizontal line is short, but not too short. The FLOORS and CEILINGS are considered delimiters and stretch to fit their argument, just as square brackets, parentheses, curly brackets, and similar characters do. Like the other delimiters, they are aligned on the math centerline. In mathematics, it is not required that these always pair left/right: some notation pairs a floor with a ceiling character. See Figure 45.

The ideally-formed half bracket is half a bracket. Comparing all of these also with square brackets: ⌞ ⌜ ⌊[ˌ⌈[X]⌉ˌ]⌋ ⌝ ⌟.

In order to distinguish FLOORS and CEILINGS properly from QUINE CORNERS, it is important that only corners may be equal sided and that FLOORS and CEILINGS are tall with short horizontals.

Scholars from non-mathematical disciplines have complained that the presentation behaviour of the mathematical characters does not satisfy their need for true ⌜half brackets⌝ with ˌpunctuation shapesˌ related to those of [square brackets] and properties. Some scholars use the mathematical shapes, which is certainly up to them. But because the mathematics community does not want to see HALF-BRACKET shapes applied to the QUINE CORNERS or the FLOORS and CEILINGS, those scholars who do prefer such shapes have no characters to use unless the HALF BRACKETS as proposed here are accepted.

4. Review of Avestan punctuation. A set of generic characters are required to support text in Avestan, Pahlavi, Manichaean, and Latin- and Arabic-script editorial practice regarding these scripts. Avestan punctuation displays the usual inconsistency of any early writing system. The punctuation proposed here for encoding is based on the system Geldner established in his 1880 edition of the *Avesta*. The ՟ AVESTAN SEPARATION POINT follows each word in Avestan and is usually followed by a space (as is most Avestan punctuation); it is *not* a sentence delimiter like U+002E FULL STOP, and indeed Geldner uses it alongside the FULL STOP in his edition (see Figure 15). The ՟ AVESTAN COLON, ՟ AVESTAN SEMICOLON, and ՟ AVESTAN END OF SENTENCE indicate breaks of increasing finality, followed by ՟ AVESTAN END OF SECTION which itself may be doubled ՟՟ (sometimes with a space between ՟ ՟) for extra finality (see Figures 16 and 17). The ՟ AVESTAN TURNED END OF SENTENCE and ՟ AVESTAN TURNED END OF SECTION are also attested, but were not used by Geldner. Sometimes the AVESTAN SEPARATION POINT may precede, or follow, other punctuation. Geldner says that the ՟ AVESTAN ABBREVIATION MARK "is employed for a sign of abbreviation and repetition" and indeed he uses it extensively in his footnotes "for the sake of economizing space".

4.1. Avestan punctuation in the context of "generic" punctuation. Comparing the Avestan punctuation to the classification of generic punctuation described above, what we find is that the Avestan punctuation is based on a different typographic model. The ՟ AVESTAN SEPARATION POINT sits on the baseline as does U+002E FULL STOP, and the ՟ AVESTAN COLON, ՟ AVESTAN SEMICOLON, and ՟ AVESTAN END OF SENTENCE are *centred* on the baseline, not resting on it as does the COLON-based punctuation of the "generic" tradition. Geldner uses the Avestan punctuation in-line in English, and it can be seen that the ՟ AVESTAN END OF SENTENCE hangs from colon-height, descends below the baseline, and has larger dots than the colon. It is our recommendation, therefore, that the Avestan punctuation be disunified from similar triangular dots used in medievalist and similar traditions. Since Avestan punctuation may be used with other Iranian scripts (it is certainly used alongside Pahlavi), and in Latin- and Arabic-script discussion of Avestan, or in transliterated editions, it is appropriate *not* to use the word "Avestan" in the character names for a number of them. The following names are proposed:

.	AVESTAN SEPARATION POINT (Figures 14, 15)
՟	AVESTAN ABBREVIATION MARK (Figure 13)
՟	RING POINT (Figure 46)
՟	TINY TWO DOTS OVER ONE DOT PUNCTUATION (Figure 17)
՟	SMALL TWO DOTS OVER ONE DOT PUNCTUATION (Figures 16, 17)
՟	LARGE TWO DOTS OVER ONE DOT PUNCTUATION (Figures 13, 14, 16, 17)
՟	LARGE ONE DOT OVER TWO DOTS PUNCTUATION (Figure 14)
՟	LARGE TWO RINGS OVER ONE RING PUNCTUATION (Figures 13, 17)
՟	LARGE ONE RING OVER TWO RINGS PUNCTUATION (Figure 31)

5.1 Comparison of relative glyph heights of Avestan and medievalist punctuation. Low, middle and high punctuation are shown in the first three lines, and then Avestan punctuation is compared with the generic colon-based punctuation. The baseline and x-height line are also shown, as well as Latin and Avestan letters for context. In the third line THEREFORE and BECAUSE are also shown.

5.2 Comparison of relative glyph heights of paired punctuation. Currently-encoded characters are given in red. The guidelines show the baseline and caps height.

6. Unicode Character Properties. Character properties are proposed here.

```
2E1A;HYPHEN WITH DIAERESIS;Pd;0;ON;;;;;N;;;;;
2E1B;TILDE WITH RING ABOVE;Pd;0;ON;;;;;N;;;;;
2E1E;TILDE WITH DOT ABOVE;Pd;0;ON;;;;;N;;;;;
2E1F;TILDE WITH DOT BELOW;Pd;0;ON;;;;;N;;;;;
2E20;POSITURA;Po;0;ON;;;;;N;;;;;
2E21;LOW PUNCTUS VERSUS;Po;0;ON;;;;;N;;;;;
2E22;COMMA ELEVATA;Po;0;ON;;;;;N;;;;;
2E23;MEDIEVAL COMMA;Po;0;ON;;;;;N;;;;;
2E24;PUNCTUS ELEVATUS;Po;0;ON;;;;;N;;;;;
2E25;PUNCTUS FLEXUS;Po;0;ON;;;;;N;;;;;
2E26;PUNCTUS VERSUS;Po;0;ON;;;;;N;;;;;
2E27;COLON WITH COMMA ELEVATA POSITURA;Po;0;ON;;;;;N;;;;;
2E28;COLON WITH RAISED COMMA POSITURA;Po;0;ON;;;;;N;;;;;
2E29;COLON WITH POSITURA;Po;0;ON;;;;;N;;;;;
2E2A;TWO DOTS OVER ONE DOT PUNCTUATION;Po;0;ON;;;;;N;;;;;
2E2B;TWO DOTS OVER COMMA POSITURA;Po;0;ON;;;;;N;;;;;
2E2C;ONE DOT OVER TWO DOTS PUNCTUATION;Po;0;ON;;;;;N;;;;;
2E2D;THREE DOTS WITH COMMA POSITURA;Po;0;ON;;;;;N;;;;;
2E2E;SQUARED FOUR DOT PUNCTUATION;Po;0;ON;;;;;N;;;;;
2E2F;FIVE DOT MARK;Po;0;ON;;;;;N;;;;;
2E30;SIMPLEX DUCTUS;Po;0;ON;;;;;N;;;;;
2E31;DICTINCTIO;Po;0;ON;;;;;N;;;;;
2E32;PUNCTUS EXCLAMATIVUS;Po;0;ON;;;;;N;;;;;
2E33;PUNCTUS INTERROGATIVUS;Po;0;ON;;;;;N;;;;;
2E34;PUNCTUS PERCONTATIVUS;Po;0;ON;;;;;N;;;;;
2E35;SIGNE DE RENVOI;Po;0;ON;;;;;N;;;;;
2E36;VIRGULA INTERPOLATIVA;Po;0;ON;;;;;N;;;;;
2E37;PARAGRAPHUS;Po;0;ON;;;;;N;;;;;
2E38;INTERPUNCTUM CIRCLE;Po;0;ON;;;;;N;;;;;
2E39;MIDDLE COMMA;Po;0;ON;;;;;N;;;;;
2E40;LEFT VERTICAL BAR WITH QUILL;Pi;0;ON;;;;;Y;;;;;
2E41;RIGHT VERTICAL BAR WITH QUILL;Pf;0;ON;;;;;Y;;;;;
2E42;TOP LEFT HALF BRACKET;Ps;0;ON;;;;;Y;;;;;
2E43;TOP RIGHT HALF BRACKET;Pe;0;ON;;;;;Y;;;;;
2E44;BOTTOM LEFT HALF BRACKET;Ps;0;ON;;;;;Y;;;;;
2E45;BOTTOM RIGHT HALF BRACKET;Pe;0;ON;;;;;Y;;;;;
2E46;LEFT SIDEWAYS U BRACKET;Ps;0;ON;;;;;Y;;;;;
2E47;RIGHT SIDEWAYS U BRACKET;Pe;0;ON;;;;;Y;;;;;
2E48;LEFT DOUBLE PARENTHESIS;Ps;0;ON;;;;;Y;;;;;
2E49;RIGHT DOUBLE PARENTHESIS;Pe;0;ON;;;;;Y;;;;;

10B37;AVESTAN SEPARATION POINT;Po;0;ON;;;;;N;;;;;
10B38;AVESTAN ABBREVIATION MARK;Po;0;ON;;;;;N;;;;;
10B39;RING POINT;Po;0;ON;;;;;N;;;;;
10B3A;TINY TWO DOTS OVER ONE DOT PUNCTUATION;Po;0;ON;;;;;N;;;;;
10B3B;SMALL TWO DOTS OVER ONE DOT PUNCTUATION;Po;0;ON;;;;;N;;;;;
10B3C;LARGE TWO DOTS OVER ONE DOT PUNCTUATION;Po;0;ON;;;;;N;;;;;
10B3D;LARGE ONE DOT OVER TWO DOTS PUNCTUATION;Po;0;ON;;;;;N;;;;;
10B3E;LARGE TWO RINGS OVER ONE RING PUNCTUATION;Po;0;ON;;;;;N;;;;;
10B3F;LARGE ONE RING OVER TWO RINGS PUNCTUATION;Po;0;ON;;;;;N;;;;;
```

7. Bibliography.

Bartoli, Francesco de. 1470. *Historia quomodo beatus Franciscus petivit a Christo indulgentiam pro Ecclesia S. Mariae de Angelis*. Trevi: Johannes Rothmann.

Björn Ellertsson. 1993. *Íslensk Þýsk orðabók: Isländisch-deutsches Wörterbuch*. Reykjavík: Iðunn. ISBN 9979-1-0220-9

Brown, Michelle P. 1993. *A Guide to Western Historical Scripts from Antiquity to 1600*. London: The British Library.

Cajori, Florian. 1952. *A history of mathematical notations*. *Volume I: Notations in elementary mathematics; Volume II: Notations mainly in higher mathematics*. La Salle and Chicago: Open Court.

Cleonardo, Nicolao. 1589. *Tabula in grammaticen Hebraeam auctore Nicolao Clenardo*. Lugduni Batavorum: Ex Officina Plantiniana, Apud Franciscum Raphelengium.

Corpus Inscriptionum Latinarum. Vol. II editio altera. Inscriptiones Hispaniae Latinae. Pars XIV Conventus Tarraconensis. Fasc. I Pars meridionalis conventus Tarraconensis. Edid. G. Alfödy, M. Clauss, M. Mayer Olivá adiuvantibus J. Corell Vicent, F. Beltrán Lloris, G. Fabre, F. Marco Simón, I. Rodà de Llanza. Berolin, Novi Eboraci 1995 ISBN 3-11-014304-6

Farley, A. (Ed.). 1783. *Domesday Book: seu liber censualis Wilhelmi primi Regis Angliæ, inter archivos regni in domo capitulari Westmonasterii asservatus: jubente rege … Georgio Tertio prælo mandatus typis.* [London].

Geldner, Karl F. 1880. *Avesta: the sacred books of the Parsis.* Stuttgart: W. Kohlhammer. Reprinted in 2003 with an introduction in Persian by Dr Jaleh Amouzgar.

Guðvarður Már Gunnlaugsson, ed. 2001. *Konungsbók Eddukvæða. Codex Regius. Stofnun Árna Magnússonar á Íslandi. Gl. Kgl. Sml. 2365 4to.* (Íslensk miðaldahandrit, 3) Reykjavík: Lögberg. ISBN: 997932161-X.

Haugen, Odd Einar. 1992. *Stamtre og tekstlandskap. Studiar i resensjonsmetodikk med grunnlag i Niðrstigningar saga.* 2 vols. Dr. philos. dissertation. Department of Scandinavian languages and literature, University of Bergen.

Haugen, Odd Einar, ed. 2004. *Handbok i norrøn filologi.* Bergen: Fagbokforlaget. ISBN 82-450-0105-8

Holm-Olsen, Ludvig, ed., 1945. *Konungs skuggsjá.* Gammelnorske tekster, 1. Oslo: Norsk Historisk Kjeldeskrift-Institutt.

Loew, E A. 1914. *The Beneventan Script. A history of the South Italian minuscule.* 1999 special edition. London: Clarendon Press ISBN 0-19-924015-9.

Núñez Contreras, Luis. *Manual de paleografía: fundamentos e historia de la escritura latina hasta el siglo VIII.* Madrid: Cátedra. ISBN 84-376-1245-4.

Ouy, Gilbert. 1987. "Ortographe et ponctuation dans les manuscrits autographes des humanistes français des XIVe et XVe siècles", in Alfonso Maierù (ed.) *Grafia e interpunzione del Latino nel medioevo, Seminario Internazionale (Roma, 27-29 settembre 1984).* Roma: Edizioni dell'Ateneo. p. 167–206.

Parkes, M. B. 1993. *Pause and effect: an introduction to the history of punctuation in the West.* Berkeley & Los Angeles: University of California Press. ISBN 0-520-07941-8.

Poure-Davoud, Ebrahim. 2001 (=1380 AP) [reprint]. *Yasna, a part of the Avesta.* Tehran: Asatir. ISBN 964-331-040-X.

Reichelt, Hans. 2004 (1383 AP). *Avesta reader: an approach to the Zoroaster's Gathas and new Avestan texts.* Translated and annotated with Persian translation of hymns and texts by Jalil Doostkhah. Tehran: Qoqnoos Publishing. ISBN 964-311-473-2.

Tertullian, Quintus Septimus Florens. [1493]. *Apologeticus adversus gentes.* Venetiis: B. Benalius.

Thompson, Edward Maunde. 1912. *An introduction to Greek and Latin palaeography.* Oxford: Clarendon Press. ISBN 0-19-925650-0.

Virgile. 1509. *Opera com. de Servius.* Milano: Leonardo Pachel.

Wahrig, Gerhard. 1974. *Deutsches Wörterbuch.* Bertelsmann Lexikon-Verlag. ISBN 3-570-06588-X

West, Martin L. 1973. *Textual criticism and editorial technique applicable to Greek and Latin texts.* Stuttgart: B. G. Teubner. ISBN 3-519-07402-8

Хабургаев, Г. А. 1986. *Старославянский язык: издание второе, переработанное и дополненное.* Москва: Просвещение.

Acknowledgements

TABLE xx - Row 2E: SUPPLEMENTARY PUNCTUATION

	2E0	2E1	2E2	2E3	2E4	2E5	2E6	2E7
0	⌜	⟋	ר	7	⊦			
1	⌜˙	⟍	⸢	·	⊣			
2	⸢	˒	◡	⸢	⌐			
3	⸣	⸓	⸣	ⸯ	⌐			
4	⸢˙	⸗	⸤	⸮	⌐			
5	⸣˙	⸗	⸥	⸪	⌐			
6	⊤	⸖	⸦	⸍	⊂			
7	⊤˙	⸗	⸨	⸧	⊃			
8	⸠	⸡	⸩	○	⸨			
9	⸟	⸰	⸪	،	⸩			
A	⸞	⸙	⸫					
B	⧠	⸲	⸬					
C	⟍	⟍	⸭					
D	⟋	⟋	⸮					
E	⸗	ⸯ	⸬					
F	▁	ⸯ	⸭					

G = 00
P = 00

9

TABLE XXX - Row 2E: SUPPLEMENTARY PUNCTUATION

hex	Name	hex	Name
00	RIGHT ANGLE SUBSTITUTION MARKER	59	(This position shall not be used)
01	RIGHT ANGLE DOTTED SUBSTITUTION MARKER	5A	(This position shall not be used)
02	LEFT SUBSTITUTION BRACKET	5B	(This position shall not be used)
03	RIGHT SUBSTITUTION BRACKET	5C	(This position shall not be used)
04	LEFT DOTTED SUBSTITUTION BRACKET	5D	(This position shall not be used)
05	RIGHT DOTTED SUBSTITUTION BRACKET	5E	(This position shall not be used)
06	RAISED INTERPOLATION MARKER	5F	(This position shall not be used)
07	RAISED DOTTED INTERPOLATION MARKER	60	(This position shall not be used)
08	DOTTED TRANSPOSITION MARKER	61	(This position shall not be used)
09	LEFT TRANSPOSITION BRACKET	62	(This position shall not be used)
0A	RIGHT TRANSPOSITION BRACKET	63	(This position shall not be used)
0B	RAISED SQUARE	64	(This position shall not be used)
0C	LEFT RAISED OMISSION BRACKET	65	(This position shall not be used)
0D	RIGHT RAISED OMISSION BRACKET	66	(This position shall not be used)
0E	EDITORIAL CORONIS	67	(This position shall not be used)
0F	PARAGRAPHOS	68	(This position shall not be used)
10	FORKED PARAGRAPHOS	69	(This position shall not be used)
11	REVERSED FORKED PARAGRAPHOS	6A	(This position shall not be used)
12	HYPODIASTOLE	6B	(This position shall not be used)
13	DOTTED OBELOS	6C	(This position shall not be used)
14	DOWNWARDS ANCORA	6D	(This position shall not be used)
15	UPWARDS ANCORA	6E	(This position shall not be used)
16	DOTTED RIGHT-POINTING ANGLE	6F	(This position shall not be used)
17	DOUBLE OBLIQUE HYPHEN	70	(This position shall not be used)
18	*INVERTED INTERROBANG*	71	(This position shall not be used)
19	*PALM BRANCH*	72	(This position shall not be used)
1A	HYPHEN WITH DIAERESIS	73	(This position shall not be used)
1B	TILDE WITH RING ABOVE	74	(This position shall not be used)
1C	LEFT LOW PARAPHRASE BRACKET	75	(This position shall not be used)
1D	RIGHT LOW PARAPHRASE BRACKET	76	(This position shall not be used)
1E	TILDE WITH DOT ABOVE	77	(This position shall not be used)
1F	TILDE WITH DOT BELOW	78	(This position shall not be used)
20	POSITURA	79	(This position shall not be used)
21	LOW PUNCTUS VERSUS	7A	(This position shall not be used)
22	COMMA ELEVATA	7B	(This position shall not be used)
23	MEDIEVAL COMMA	7C	(This position shall not be used)
24	PUNCTUS ELEVATUS	7D	(This position shall not be used)
25	PUNCTUS FLEXUS	7E	(This position shall not be used)
26	PUNCTUS VERSUS	7F	(This position shall not be used)
27	COLON WITH COMMA ELEVATA POSITURA		
28	COLON WITH RAISED COMMA POSITURA		
29	COLON WITH POSITURA		
2A	TWO DOTS OVER ONE DOT PUNCTUATION		
2B	TWO DOTS OVER COMMA POSITURA		
2C	ONE DOT OVER TWO DOTS PUNCTUATION		
2D	THREE DOTS WITH COMMA POSITURA		
2E	SQUARED FOUR DOT PUNCTUATION		
2F	FIVE DOT MARK		
30	SIMPLEX DUCTUS		
31	DISTINCTIO		
32	PUNCTUS EXCLAMATIVUS		
33	PUNCTUS INTERROGATIVUS		
34	PUNCTUS PERCONTATIVUS		
35	SIGNE DE RENVOI		
36	VIRGULA INTERPOLATIVA		
37	PARAGRAPHUS		
38	INTERPUNCTUM CIRCLE		
39	MIDDLE COMMA		
3A	(This position shall not be used)		
3B	(This position shall not be used)		
3C	(This position shall not be used)		
3D	(This position shall not be used)		
3E	(This position shall not be used)		
3F	(This position shall not be used)		
40	LEFT VERTICAL BAR WITH QUILL		
41	RIGHT VERTICAL BAR WITH QUILL		
42	TOP LEFT HALF BRACKET		
43	TOP RIGHT HALF BRACKET		
44	BOTTOM LEFT HALF BRACKET		
45	BOTTOM RIGHT HALF BRACKET		
46	LEFT SIDEWAYS U BRACKET		
47	RIGHT SIDEWAYS U BRACKET		
48	LEFT DOUBLE PARENTHESIS		
49	RIGHT DOUBLE PARENTHESIS		
4A	(This position shall not be used)		
4B	(This position shall not be used)		
4C	(This position shall not be used)		
4D	(This position shall not be used)		
4E	(This position shall not be used)		
4F	(This position shall not be used)		
50	(This position shall not be used)		
51	(This position shall not be used)		
52	(This position shall not be used)		
53	(This position shall not be used)		
54	(This position shall not be used)		
55	(This position shall not be used)		
56	(This position shall not be used)		
57	(This position shall not be used)		
58	(This position shall not be used)		

TABLE XXX - Row 10B: AVESTAN

	10B0	10B1	10B2	10B3
0	ﺎﻠ	ﻭ	ﺮﻟ	ﻊ
1	ﺳﻟ	ﻥﻉ	ﺍﻟﻟ	ﻊﻌ
2	ﺲﻌ	ﺶﻌ	٣	ﺍﻟﻟ
3	ﺲﻌﻌ	ﺶﻌ	ﻊﻌ	ﻊﻌﻌ
4	ﺥﻉ	ﻉﻉ	٣	ﺍﻟﻟ
5	ﻙﻉ	ﻥﻥ	١	ﺍﻟﻟ
6	٤	ﻥﻉ	ﻝ	
7	٤	ﻡ	ﻙﻉ	.
8	ﻡﻟ	ﻙﻉ	ﻑ	○
9	ﻙﻟ	ﻡﻉ	ﻑ	○
A	ﺪ	ﻉ	ﻙ	∴
B	ﻑ	ﻭ	ﻡﻟ	∴
C	،	ﻙ	ﻙ	•• •
D	ﺪ	ﻕﻉ	ﺍ	•• ••
E	ﻭ	ﻭﻟ	ﻑ	○ ○ ○
F	ﺪ	ﻭﻟ	ﻥﻥ	○ ○ ○

G = 00
P = 01

TABLE XXX - Row 10B: AVESTAN

hex	Name		hex	Name
00	AVESTAN LETTER A			
01	AVESTAN LETTER AA			
02	AVESTAN LETTER AO			
03	AVESTAN LETTER AAO			
04	AVESTAN LETTER AN			
05	AVESTAN LETTER AAN			
06	AVESTAN LETTER AE			
07	AVESTAN LETTER AEE			
08	AVESTAN LETTER E			
09	AVESTAN LETTER EE			
0A	AVESTAN LETTER O			
0B	AVESTAN LETTER OO			
0C	AVESTAN LETTER I			
0D	AVESTAN LETTER II			
0E	AVESTAN LETTER U			
0F	AVESTAN LETTER UU			
10	AVESTAN LETTER KE			
11	AVESTAN LETTER XE			
12	AVESTAN LETTER XYE			
13	AVESTAN LETTER XVE			
14	AVESTAN LETTER GE			
15	AVESTAN LETTER GGE			
16	AVESTAN LETTER GHE			
17	AVESTAN LETTER CE			
18	AVESTAN LETTER JE			
19	AVESTAN LETTER TE			
1A	AVESTAN LETTER THE			
1B	AVESTAN LETTER DE			
1C	AVESTAN LETTER DHE			
1D	AVESTAN LETTER TTE			
1E	AVESTAN LETTER PE			
1F	AVESTAN LETTER FE			
20	AVESTAN LETTER BE			
21	AVESTAN LETTER BHE			
22	AVESTAN LETTER NGE			
23	AVESTAN LETTER NGYE			
24	AVESTAN LETTER NGVE			
25	AVESTAN LETTER NE			
26	AVESTAN LETTER NYE			
27	AVESTAN LETTER NNE			
28	AVESTAN LETTER ME			
29	AVESTAN LETTER HME			
2A	AVESTAN LETTER YYE			
2B	AVESTAN LETTER YE			
2C	AVESTAN LETTER VE			
2D	AVESTAN LETTER RE			
2E	AVESTAN LETTER LE			
2F	AVESTAN LETTER SE			
30	AVESTAN LETTER ZE			
31	AVESTAN LETTER SHE			
32	AVESTAN LETTER ZHE			
33	AVESTAN LETTER SHYE			
34	AVESTAN LETTER SSHE			
35	AVESTAN LETTER HE			
36	(This position shall not be used)			
37	AVESTAN SEPARATION POINT			
38	AVESTAN ABBREVIATION MARK			
39	RING POINT			
3A	TINY TWO DOTS OVER ONE DOT PUNCTUATION			
3B	SMALL TWO DOTS OVER ONE DOT PUNCTUATION			
3C	LARGE TWO DOTS OVER ONE DOT PUNCTUATION			
3D	LARGE ONE DOT OVER TWO DOTS PUNCTUATION			
3E	LARGE TWO RINGS OVER ONE RING PUNCTUATION			
3F	LARGE ONE RING OVER TWO RINGS PUNCTUATION			

Examples

Figure 1. Sample from Holm-Olsen 1945 showing LEFT VERTICAL BAR WITH QUILL, and RIGHT VERTICAL BAR WITH QUILL.

Figure 2. Sample from Farley 1783, showing PUNCTUS ELEVATUS.

Figure 3. Sample from Farley 1783, showing PARAGRAPHUS.

Figure 4. Sample from Cleonardo 1589, showing TWO DOTS OVER ONE DOT PUNCTUATION, here used to educate the reader as to the shape of a HEBREW POINT SEGOL.

Figure 5. Sample from Virgile 1509, showing PARAGRAPHUS.

Figure 6. Sample from Tertullian [1493], showing PUNCTUS ELEVATUS.

Figure 7. Sample from Bartoli 1470, showing SQUARED FOUR DOT PUNCTUATION.

Teikn i tekstutgåver

Teikna skifter noko frå utgåve til utgåve, men mange norrøne tekstutgåver held seg
til dette utvalet:

I teksten

… ǀ …	linjeskifte
… ǁ …	spalte- eller sideskifte
⟨ … ⟩	tekst tilføydd av utgjevar (ofte for å gjere lesinga lettare eller setninga korrekt etter vanlege grammatiske reglar)
ⱶ … ⱶ	ekspunksjon, dvs. overstreka, underprikka eller utradert i handskriftet
ⱶ … ⱶ	dittografi, dvs. same ord to gonger på rad, utan at det er retta i hand-skriftet
[…]	uleseleg i handskriftet, med framlegg til utfylling av utgjevaren
` … ´	tilføydd over linja, anten av skrivaren sjølv eller med yngre hand (*m.y.h.*)
⁄ … ⁀	tilføydd i margen, anten av skrivaren sjølv eller med yngre hand
ooooo	uleselege teikn, utan framlegg til utfylling av utgjevaren (like mange nullar som det er uleselege teikn; om det er mange teikn, kan det stå eit tal for kor mange teikn som truleg har stått der)
. . .	usikker lesemåte, markert med underprikking for kvart teikn det gjeld
*	merknad knytt til ordet, forklart i det kritiske apparatet
h*er*	kursivering, brukt for å vise oppløyst forkortingsteikn
O(lafr)	parentes, særleg brukt for å vise utfylling av suspensjonar (typisk markert med punktum eller anna generelt forkortingsteikn)

I det kritiske apparatet

…]	lemma, dvs. lesemåte i teksten, avslutta med hakeparentes (som dermed fungerer som eit kolon)
+ …	tekst som kjem i tillegg i eitt eller fleire andre handskrifter
÷ …	tekst som manglar i eitt eller fleire andre handskrifter
sål.	slik, dvs. for å vise at den lesemåten som er oppgitt i apparatet, førekjem i denne forma i det eller dei handskriftene som følgjer

Figure 8. Sample from Haugen 2004, showing LEFT VERTICAL BAR WITH QUILL and RIGHT VERTICAL BAR WITH QUILL.

ſtaⱱþo*m* heıma varg lıoþo*m* vanr a vıþo*m* vtı. komo þ*er* ogogn oll
at heⱨðı þa e*r* broþɜ þıno*m* brıoſt raⱱɟaþ*ır*. Goɜþ*ır* þıc ɟręgıan aɟ
3 ɟırın v*er*co*m* þv vart brvþ*r* grana a bra vellı gvll bıtlvþ vart
goɜ t*ıl* raſar. haɟða ec þ*er* moþrı mart ſceıþ rıþıt ſvangr*ı* vⱨð
ſaⱱþlı ſımvl ɟoɜbergıſ. Svєιɴ þott*ır* þv ſıþ laⱱs v*er*a þa er
6 þv gvllnıſ geıtr molcaþ*ır* єɴ ıaɴat ſıɴ ımþar ðott*ır* taⱱttrⱱg
hypıa vıll þv taⱱlo lengrı. Ɵyʀ vılða ec at ɟreca ſteını hr*aⱱ*
na ſeþıa ahręom þıno*m*. єɴ tıcr ẏþrár teygıa at ſollı eⱱþa ge
9 ɟa gaⱱlto*m* ðeılı graⱱm vıⱱþ þıc. Uęrı ycr ſınɟıotlı .q. ɟęmra
myclo gvɴı at heẏıa *oc*glaþa oɜno. єɴ ſe onyto*m* oɜþo*m* at bregðaz
þot hr*ı*ng br*o*tar heıpt*ır* ðeılı. Þıccıat m*er* goð*ır* granmarſ ſyn*ır*
12 þo ðvg*ır* ſıclıngo*m* ſatt at męla. þeır haɟa marcat a móıⱨs heımo*m*
at hvg haɟa hıoɜıo*m* at bregða. Þeır aɟ rícı rєɴa leto ſvı
pvð *oc* ſvegıoþ ſol heıma t*ıl*. ðala ðaⱱɢótta ðaⱱcqⱱar hlıþ*ır* ſkalⱱ
15 mıſtar maʀ hvár meg*ır* ɟoɜo. Mętto þeır tıɢıa ıtvn hlıþı
ſaⱱgðo ſtrıþlıga ſtılı qⱱomo. vtı ſtoþ haⱱþbrodoɜ hıalmı ɟal
ðın hvgðı h*aɴ* ıoɜeıþ ettar ſınar hvı er h*er*mþar lıtr a hnıɟlvⱨ-
18 go*m*. Snvaz h*er* at ſanðı ſneɟgır kıólar racca hırt*ır* *oc* rár lan
gar. ſcılð*ır* marg*ır* ſcaɟnar árar gaⱱɟvct lıþ gylɟa glaþ*ır* ylⱱı-
ngar. Ganga xv. ɟolc vp alanð þo er ıſogn v́t vıı. þvſvn
21 ðır. lıɢıa h*er* ıgrınðo*m* ɟyr gnıpa lvnðı brım ðýr blá ſvóɜt *oc* bvıⱨ
gvllı. þar er mıclo meſt mengı þeıra mvna nv helgı hıoɜþıng
ðvⱱlaɜ reⱨı raⱱcn bıtlvþ t*ıl* regın þınga єɴ ſpóɜ vıtn*ır* at
24 ſparıⱨ heıðı. Meln*ır* *oc* myln*ır* t*ıl*. myrc vıþar latıþ engı Ⱡɤaɴ

Figure 9. Sample from Guðvarður Már Gunnlaugsson 2001, showing MEDIEVAL COMMA.

GUIDE TO HISTORICAL SCRIPTS

'Scriptorium' system, devised for use in the journal *Scriptorium*. This conflation was produced (unpublished) by Julian Brown, but is modified here.

Expanded abbreviations: expansion in round brackets. For example, notu(m) for *notū*, (et) for &. On the rare occasions where brackets are actually used in the text, entailing confusion, this is commented upon in a footnote. Another common form of expansion is by italicization, but this may be open to confusion with the practice of italicization of rubrics and titles.
Unexpanded abbreviations: the apostrophe.
Where the transcriber cannot or will not expand or where the original abbreviation is otiose.
Textual omission by the scribe: angled brackets or half-square brackets.
Empty where the omission is not made good, ⟨ ⟩. Occupied where it is made good, ⟨est⟩.
Where another source has been used and the supplement is not conjectural, half brackets may be used, ⌊est⌋.
Textual interpolation by the scribe: braces.
Where the transcriber wishes to cancel, {est}.
Problematical readings: obeli, alias daggers.
Used where the text is corrupt or obscure, or where the transcriber is unsure of the correct reading, †est†.
Scribal insertions: slashes.
On the writing line, / est \; between the lines, \ est /; marginal, \\ est //.
Cancellations: square brackets.
By washing, scraping or pouncing, cancelled letters illegible [ ], or legible [est]. By crossing out, illegible [−], legible [− est]. By expunctuation (i.e. a point placed beneath the letter to be cancelled), [ẹṣṭ].
Substitutions: square brackets and a slash.
Actually on letters cancelled by washing, scraping or pouncing, illegible cancellation, [/ est], legible cancellation, [et / est]. Above letters cancelled by crossing out, illegible cancellation, [−\ est], legible cancellation, [− et \ est]. Above letters cancelled by expunctuation, [ẹṭ \ est]. By transformation (where the original letter is adapted to form another letter), [o ⟩ a]. By simple addition (where the new letter is simply written over or above the original letter), on a suppressed letter, [o + a], above a suppressed letter, [o +\ a].
Accidental loss (trimming, rodent activity, staining, etc): double square brackets.
Number of lost letters unknown, [[ ]]. Number of lost letters estimated [[∗∗∗]] or [[3]], or approximately estimated, [[± 10]].
Letters doubtful or illegible owing to damage: asterisk.
On the line for illegible letters, quod ∗∗∗ demonstrandum. Below the line for doubtful letters, quod e̤ṣ̤t̤ demonstrandum.
Ends of manuscript lines: vertical bar, |.
Rubrics, titles, lemmata, etc.: italicization.
Punctuation and capitalization: modern pauses of equivalent or near equivalent values are substituted

Figure 10. From Brown 1993, showing BOTTOM LEFT HALF BRACKET and BOTTOM RIGHT HALF BRACKET, as well as MATHEMATICAL LEFT ANGLE BRACKET and MATHEMATICAL RIGHT ANGLE BRACKET.

⟦ ⟧Double brackets enclose letters or words that a scribe has deleted in the manuscript itself. If such letters can no longer be read, use dots as above, ⟦. . .⟧. The symbol ‖‖, repeated for each letter, is also used for successful erasures.

‘   ’ These signs are used by papyrologists to enclose insertions by a scribe after he has made his original line. One could represent a scribal alteration of δε to τε by ⟦δ⟧‘τ′ε, but it is more elegant to print simply τε and note in the apparatus ‘δε ante corr.’; if the reader fails to consult the apparatus, that is his fault.

⌊ ⌋Half brackets are a logical modification of full square brackets. In papyrus texts they indicate that the papyrus itself is broken or worn away but that the supplement is supplied by another source and is not conjectural. In other texts, by extension, they can be used to show the absence of a particular source, whether because it is damaged or because it has a shortened version of the text. They might well be used, for instance, in a text of Nonnus' Dionysiaca to show which letters are preserved in the Berlin papyrus, or in one of Athenaeus to show how much is attested by the epitome. Complications, arise, however, if there is more than one of these intermittent sources ⌈ ⌉ can be used for a second one, and Bethe's Pollux shows how more elaborate information can be conveyed; but it may be wondered whether such feats of typography are often worth while, especially when they endanger the legibility of the text.

†   † Obeli mark words which the editor judges to be corrupt. If only one word is suspect, only one obelus is needed: *subsidiis magnis †epicuri constabilitas.* If the editor cannot limit the corruption to one word, he places his two obeli so as to define the area within which it is to be sought: *declinare quis est qui †possit cernere sese†.*

Figure 11. From West 1973, describing the use of BOTTOM LEFT HALF BRACKET, BOTTOM RIGHT HALF BRACKET, TOP LEFT HALF BRACKET, and TOP RIGHT HALF BRACKET. Also shown is MATHEMATICAL LEFT WHITE SQUARE BRACKET and MATHEMATICAL RIGHT WHITE SQUARE BRACKET

EXPLICATIO NOTARVM

Et notae quae sequuntur et litterae eis e. g. additae cursivae q. d. redditae sunt notis ⊂ ⊃ exceptis, quae semper directae ponuntur. Animadverte variis locis, i. e. in exemplo apparatu critico commentario, scripturae quidem qualitatem et notarum et litterarum, quibus notae adiectae sunt, variare aptam consentaneam ei, quod editor intendit, ut perspiceretur. Unde invenies notas directas et cursivas litterasque et maiusculas et minusculas et directas et cursivas.

Versuum paginarumque divisio

ab|c, abc |ˢ abc   Versuum divisio (in apparatu critico et in falsis solis)

abc || abc   Paginarum vel laterum monumenti divisio; litterarum quae non suo versui inscriptae sunt divisio

Litterae communes

abc   Litterae planae lectionis ad verbum aut ad partem, quod quidem agnosci possit, verbi pertinentes

ABC   Litterae interpretationis incertae planae prorsus lectionis, quae tamen in incerto sunt utrum ad initium verbi an ad medium an ad finem pertineant an ad verbum plus quam unum an alio quo modo intellegendae sint; praeterea saepe tituli male traditi et lectiones variae maiusculis redduntur

V   Littera *V* typo maiusculo reddita, si incertum est quo modo phonetice explicanda sit

ạbc̦   Litterae ita laesae ut certius extra tituli contextum agnosci non possint

+++   Vestigia litterarum ita corruptarum, ut nullo modo agnosci possint; pro littera quaque ponitur singula crux

<u>*abc*</u>   Litterae quae a prioribus lectae descriptae postea perierunt

Litterarum formae notabiliores

á, é, í, ó, ú   Vocales apicibus notatae

ỉ   I longa

$\overline{abc}$, *V̄*   Litterae et notae numerorum lineis supra ductis notatae

Ʉ, ł   Litterae et notae numerorum lineis directis perductis notatae

⊂*y*⊃, ⊂*bs*⊃, ⊂*ps*⊃, ⊂*v*⊃   Litterae quae dicuntur Claudianae

ab̂, ab̂ĉ   Litterae conexae (unaquaeque littera, quae cum sequenti coniuncta est nexu, notatur accentu circumflexo superposito)

Litterae erasae

⟦*ȧḃċ*⟧, ⟦*ȦḂĊ*⟧, ⟦*ȧḃċ̦*⟧   Litterae antiquitus erasae, quae nihilominus aut plane aut probabiliter leguntur

⟦<u>*abc*</u>⟧   Litterae antiquitus erasae quas editor supplevit

⟦*[...]*⟧, ⟦*[-5?-]*⟧, ⟦*[---]*⟧   Litterae antiquitus erasae, quarum numerus aut certius aut minus certe aut nullo modo computari potest

⟦*[------]*⟧   Lacuna unius integri versus erasi

⟦*[------]*⟧   Lacuna integrorum versuum erasorum quorum numerus incertus est

Litterae in litura scriptae

⟪*abc*⟫, ⟪*ABC*⟫, ⟪*ạḇc̦*⟫   Litterae loco alterarum erasarum insculptae quae aut plane aut probabiliter leguntur

⟪*[abc]*⟫   Litterae loco alterarum erasarum insculptae, quae tamen fortuito superficiei damno aut fractura omnino perierunt quasque editor supplevit

⟪*[...]*⟫, ⟪*[-5?-]*⟫, ⟪*[---]*⟫   Litterae loco alterarum erasarum insculptae, quae tamen fortuito superficiei damno aut fractura omnino perierunt quarumque numerus aut certius aut minus certe aut nullo modo computari potest

Additamenta antiqua

`` `abc´ ``   Additamentum antiquum ad textum corrigendum aut supplendum

Figure 12a. From von Hübner 1957, showing LEFT SIDEWAYS U BRACKET and RIGHT SIDEWAYS U BRACKET. Also shown is MATHEMATICAL LEFT WHITE SQUARE BRACKET and MATHEMATICAL RIGHT WHITE SQUARE BRACKET, and MATHEMATICAL LEFT DOUBLE ANGLE BRACKET and MATHEMATICAL RIGHT DOUBLE ANGLE BRACKET.

EXPLICATIO NOTARVM   **XXX**   HISPANIA CITERIOR

Litterae deperditae

[abc]   Litterae fortuito superficiei damno aut fractura omnino deperditae quas editor supplevit

[abc-]   Litterae deperditae et suppletae ut supra, sed forma declinationis aut coniugationis incerta

[...]   Litterae omnino deperditae ut supra, quas editor supplere non potuit quamquam numerus earum haud dubie computatur; pro littera quaque ponitur singulum punctum

[-5?-], [---]   Litterae omnino deperditae ut supra, quas editor non supplevit quarumque numerus aut coniectura tantum aut nullo modo computari potest

[-]   Praenomen non perscriptum omnino deperditum cuius litterarum numerus incertus est quodque editor supplere non potuit

[------]   Lacuna unius integri versus

⟦[------]⟧   Lacuna unius integri versus erasi

------   Lacuna integrorum versuum quorum numerus incertus est

⟦------⟧ ˙   Lacuna integrorum versuum erasorum quorum numerus incertus est

[-?], [---?], [------?], ⟦[------?]⟧, ------?, ⟦------?⟧   Incertum sintne litterae fortuito superficiei damno aut fractura deperditae

Litterae ab editore expunctae, adiectae, correctae

{abc}   Litterae errore adiectae quas editor expunxit

⟨abc⟩   Litterae errore omissae quas editor adiecit

(a)bc, a(b)c, ab(c)   Litterae e consuetudinibus quibusdam orthographicis phoneticis omissae, quas editor adiecit

ˈabcˈ   Litterae quas editor correxit

Verba breviata, notae, verba subaudita, inscriptiones non finitae

a(bc)   Verbum non perscriptum quod editor solvit

a(bc-)   Verbum non perscriptum quod editor solvit ut supra, sed forma declinationis aut coniugationis incerta

a(---)   Verbum non perscriptum quod certe solvi non potest

A(---)   Nomen (in titulo ineunte: nomen aut verbum) non perscriptum quod certe solvi non potest; nota ponitur etiam, si incertum est, utrum de verbo an de nomine agatur

⊂abc⊃   Verbum ab editore loco litterae inversae aut notae in inscriptione positae inductum

⊂Manius⊃   *M'*. i. e. *Manius*

⊂crux⊃   Imago inscriptioni inserta vel adiecta, quam editor verbo uno verbisve compluribus descripsit

(scil. abc)   Verbum in textu non scriptum, sed subauditum, quod editor addidit

⟨---⟩, ⟨------⟩   Inscriptio non ad finem perducta

Interpunctio

a∘b∘c   Interpunctio inter singulas aut litteras aut syllabas aut partes unius verbi posita

abc ∘ abc   Interpunctio inter singula verba posita

Aliae notae editoris

(!)   Nota editoris ut lector attento animo aliquid perpendat

(vac. 3), (vac. 3?), (vac.)   Tractus versus qui non inscripti fuisse videntur quorumque amplitudo, numero litterarum expressa, certius aut minus certe aut nullo modo computari potest (quae res non adnotatur nisi maioris momenti videtur)

(vac.?)   Superficies ita corrupta, ut utrum inscripta fuerit necne statui non possit

..., [...]   Verbum aut verba praetermissa ab editore in laudando aut textum antiquum aut recentem

ab[c?], ab(c?), ab{c?}, ab⟨c?⟩, abˈc?ˈ, bc (?)   Nota editoris quae ponitur post verbum aut syllabam aut litteram, si suppleta vel soluta vel expuncta vel adiecta vel correcta vel lectio interpretatiove incerta sunt

                              H. K.

Figure 12b. From von Hübner 1957, showing LEFT SIDEWAYS U BRACKET and RIGHT SIDEWAYS U BRACKET. Also shown is MATHEMATICAL LEFT WHITE SQUARE BRACKET and MATHEMATICAL RIGHT WHITE SQUARE BRACKET, and MATHEMATICAL LEFT ANGLE BRACKET and MATHEMATICAL RIGHT ANGLE BRACKET.
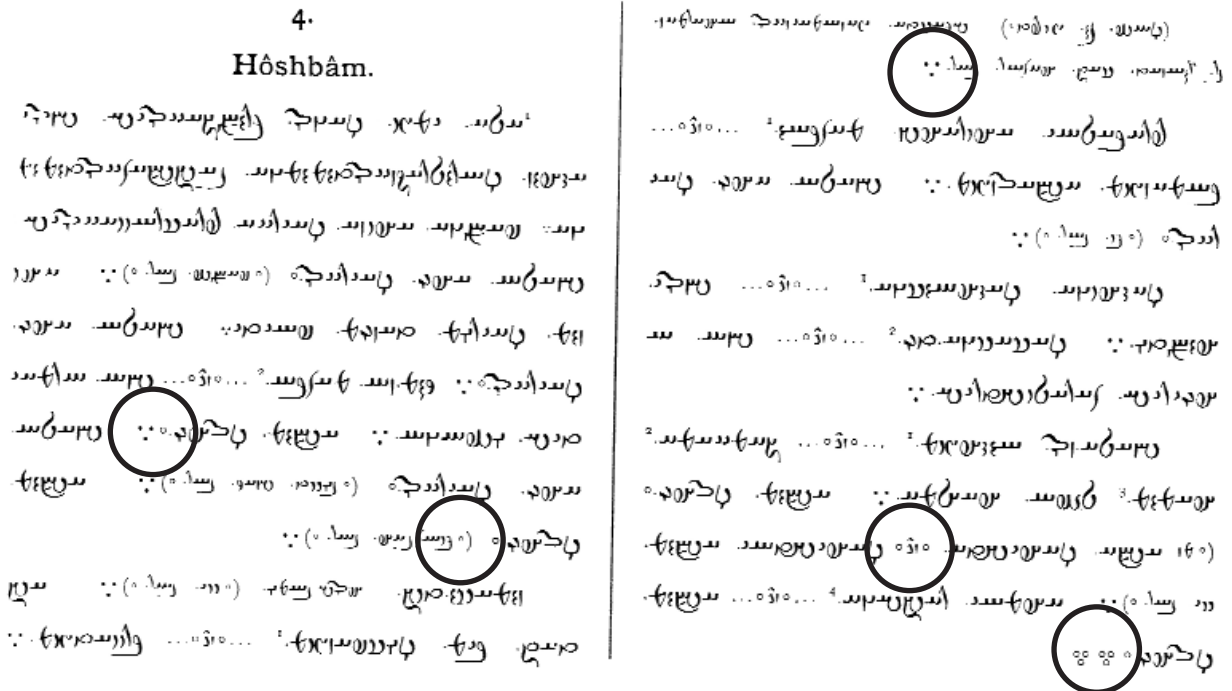
Figure 13. Avestan text from Geldner 1880, showing AVESTAN ABBREVIATION MARK, LARGE TWO DOTS OVER ONE DOT PUNCTUATION, and LARGE TWO RINGS OVER ONE RING PUNCTUATION.
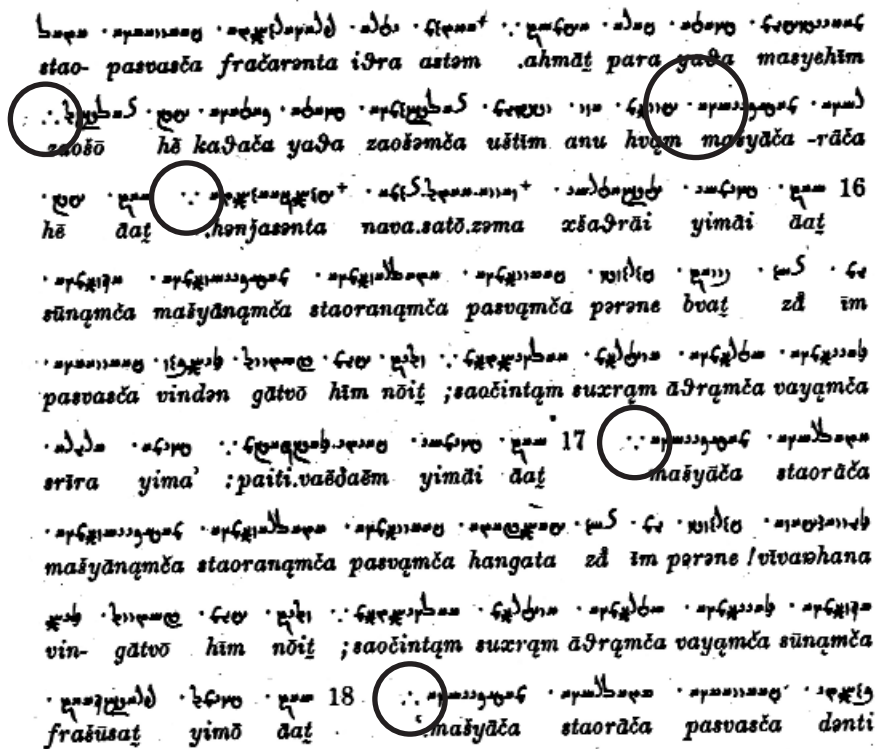
Figure 14. Avestan text from Reichelt 2004, showing AVESTAN SEPARATION MARK, LARGE TWO DOTS OVER ONE DOT PUNCTUATION, and LARGE ONE DOT OVER TWO DOTS PUNCTUATION.
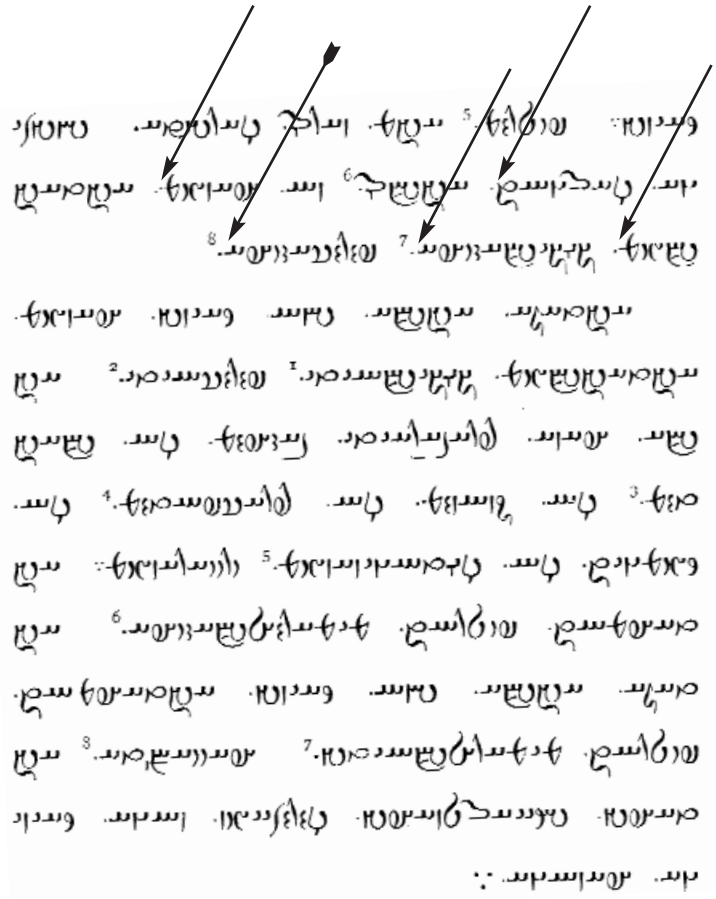
Figure 15. Passage of Avestan text from Geldner 1880 showing FULL STOP (with fletched arrow) alongside AVESTAN SEPARATION POINTS.
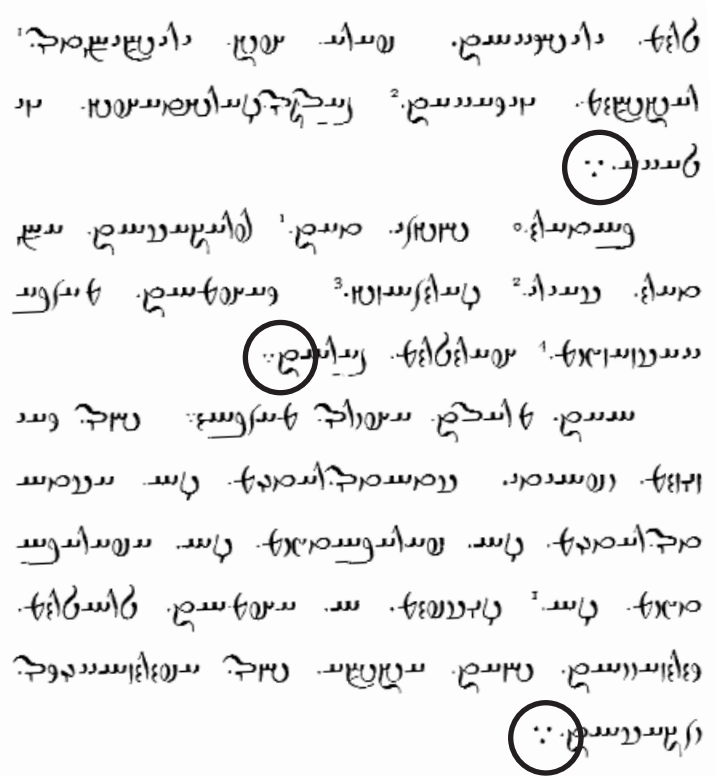
Figure 16. Passage of Avestan text from Geldner 1880 showing SMALL TWO DOTS OVER ONE DOT PUNCTUATION (centre) alongside LARGE TWO DOTS OVER ONE DOT PUNCTUATION (top and bottom).

> Punctuation is a weak point in the manuscripts. They have various signs of punctuation, indeed, but they have no system of punctuation. I had to devise a system for myself, based upon the signs that are found: thus, ⸴ is a colon, ⹁ semicolon or period, ⁖ denotes the close of a paragraph or the end of a sentence, ⸎ a greater division, ⸎⸎ the end of a chapter. In texts with Pahlaví and Sanskrit translation, this version takes the place of punctuation. The text is cut up by the translation into sentences of varied length, sometimes of only one or two words. For the most part, however, the explanation is inserted at an actual break in the sentence. It is only in the rarest instances that the manuscripts vary in this traditional division of the sentence.[1] In the sâda-texts, on the contrary, the text runs along without break until the first large division, with a mark of punctuation thrown in only here and there.[2] A general exception is made in the Gâthâs, where the close of the strophe is regularly denoted by the star. Where the translation was wanting I was therefore often thrown upon my own resources as regards punctuation. In that case the punctuation may be regarded in doubtful instances simply as subjective. In the Vendidâd I have adhered almost absolutely to the division of sentences by the Pahlaví translation, and have indicated the same either by a larger point (.), or by ⸴ where the sentence ends. I regret that I did not do this consistently also in the Yasna; although even there the larger point for the most part denotes a division by the translation.[3] Sometimes, however, I have departed from

Figure 17. Passage from Geldner's 1880 grammar discussing his systematic use of ⸴ TINY TWO DOTS OVER ONE DOT PUNCTUATION, ⸴ SMALL TWO DOTS OVER ONE DOT PUNCTUATION, ⁖ LARGE TWO DOTS OVER ONE DOT PUNCTUATION, and ⸎ LARGE TWO RINGS OVER ONE RING PUNCTUATION, as well as his use of sentence-final FULL STOP, to be distinguished from AVESTAN SEPARATION POINT.

> **§ 33.** Древнее славянское письмо было с п л о ш н ы м: слова в строке не выделялись, заглавные буквы употреблялись в начале глав, но не предложений (см. фотокопии отрывков из памятников). Конец главы обычно обозначался комбинацией точек и черточек: · ┆ ⁝ , ⸭ ╷ , : — и т. д.

Figure 18. Sample from Хабургаев 1986 showing FIVE DOT MARK.

> верема [59] Ростиславъ Смоленьскꙑ
> дчери [61] оу Стослава. оу [62] Шлгоⷩ
> сн҃а своего [63]. Смоленьскоу. и ве
> Новагорода в нѣⷧю по водохрⷭ
> геньварꙗ въ о҃ дн҃ь ⁖ в то же ве
> славъ поиде на Гюргꙗ стрꙑꙗ [67] с
> своего Володимира [68] ѿстави в Кꙑ

Figure 19. Sample from a 1908 edition of the Ипатьевская летопиць, showing ONE DOT OVER TWO DOTS PUNCTUATION.

Contritio et infelicitas in uiis eorum: et uiam pacis non
cognouerunt ." non est timor dei ante oculos eorum .

Figure 21. Sample from Parkes 1993 showing PUNCTUS FLEXUS, PUNCTUS ELEVATUS, and LOW PUNCTUS
VERSUS.

Signos de puntuación en textos españoles

Los signos de puntuación en códices y documentos españoles de la Edad
Media es muy irregular. Agustín Millares Carlo, uno de los más rigurosos
especialistas en escritura visigótica, afirmó que «los signos en códices y
documentos de escritura visigótica pertenecen a más de un sistema y su
equivalencia respecto a los actuales no se deja precisar»[27]. Los signos son
éstos:

A partir del siglo XII y en textos en escritura carolingia se empezó a
regularizar el sistema de puntuación con notable mengua de los signos: casi
en exclusiva se utilizaron para indicar las pausas el punto (.) en línea de
escritura, y el punto con vírgula sobrepuesta (/); la pausa final se indicó con
un signo formado así: ∴   ؛   ؛

En los documentos escritos en castellano en la baja Edad Media, la
puntuación es muy irregular, adoptándose rayas y curvas acompañadas de
puntos a veces para señalar y no con exactitud los periodos.

Figure 22. Sample from Núñez Contreras 1994 showing FULL STOP, PUNCTUS VERSUS, COLON,
ONE DOT OVER TWO DOTS PUNCTUATION, SQUARED FOUR DOT PUNCTUATION,
TWO DOTS OVER ONE DOT PUNCTUATION, GEORGIAN PARAGRAPH SEPARATOR, PUNCTUS ELEVATUS,
MEDIEVAL COMMA, PUNCTUS FLEXUS, SEMICOLON, PUNCTUS INTERROGATIVUS, and COLON WITH POSITURA.

lestrarbók *kv.* **-ar, -bækur**: Lesebuch *hk.* -(e)s,
    ⸗er.
lestrarfýsn *kv.* **-ar**: Leselust *kv.* -, ⸗e.
lestrarmerki *hk.* **-s, -**: Satzzeichen *hk.* -s,
lestrarsalur *kk.* **-ar, -ir**: Lesesaal *kk.* -(e)s, ⸗e.
lestur *kk.* **-urs** *eða* **-rar, rar**: Lesen *hk.* -s; Studium
    hk. -s, -ien; Pflücken *hk.* -s.
leti *kv.* **-**: Faulheit *kv.* -.
letidýr *hk.* **-s, -**: Faultier *hk.* -(e)s, -e.
letigarður *kk.* **-s, -ar**: Strafanstalt *kv.* -, -en.

Figure 23. Sample from Björn Ellertsson 1993 showing HYPHEN WITH DIAERESIS.

3 **Eine Tilde** (~) wurde verwendet, um sprachlich
  zusammengehörige Wörter miteinander zu einer
  größeren Einheit zu verbinden und dabei Raum
  einzusparen. Die Tilde ersetzt das am Anfang eines
  solchen größeren Artikels stehende Wort oder den
  durch einen senkrechten Strich (|) abgetrennten er-
  sten Teil dieses Wortes; z. B.:

  belebt . . .   ~heit (~ Belebtheit) . . .

  Augen|abstand . . . ~ader (= Augen-
  ader) . . .   **~arzt** (= Augenarzt) . . .
  ~blick (= Augenblick) . . .   **~blick-
  lich** (= augenblicklich)

  Der über oder unter einer Tilde stehende Punkt
  bedeutet dabei, daß ein Stichwort groß (~̇) oder
  klein (~̣) geschrieben wird.

Figure 24. Sample from Wahrig 1974 showing TILDE WITH DOT ABOVE and TILDE WITH DOT BELOW.

Cicero, *Cato maior de senectute*, x,32 – xi,35. (= P).

(Col. a, line 23)

Cum ingressus iter pedibus sit . in [a]equum omnino non
   ascendere ˙
Cum autem equo . ex equo non descendere ˙ Nullo imbre .
   nullo frigore adduci . ut capi|te operto sit, summam esse in
   eo corporis || (col. b) siccitatem ˙
Itaque omnia exequi regis officia & munera ˙
Potest igitur exercitatio & temperantia ˙ etiam senectute
   conseruare aliquid pristini roboris ˙
Ne desint in senectute uires . Ne postulantur quidem uires a
   senectute ˙
Ergo & legibus & institutis uacat aetas nostra . muneribus iis
   quoniam possunt sine uiribus sustineri ˙
Itaque non modo quod non possumus . sed ne quantum
   possumus quidem cogimur ˙

[that] When [Masinissa] begins a journey on foot . he does
   not mount a horse at all ˙
When on horseback. he does not dismount ˙ No rain . no
   cold can bring him . to go with head covered . in body he is
   extremely || (col. b) wiry ˙
Thus to perform all the duties of a King ˙
Therefore one can by exercise and self control . preserve even
   in old age some original vigour ˙
Lest strength fail in old age . none is even expected from old
   age ˙
Hence both by law and tradition men of our age . abandon
   those duties which can (not) be sustained without strength ˙
Therefore we are not only not required to do what we cannot
   . but we are not even compelled to do as much as we can ˙

At multi ita sunt inbecilli senes . ut nullum of|ficii aut
   omnino uitae munus exequi possint ˙
At id quidem non proprium senectutis uitium est . sed com-
   mune ualitudinis ˙
Quam fuit inbecillus P. africanus filius is qui te adoptauit .
   quam tenui aut nulla potius ual' i 'tudine.
Quod ni ita fuisset . alterum illud extitisset lumen ciuitatis ˙
Ad paternam enim magnitudinem . animi doctrina uberior
   accesserat ˙
Quid mirum igitur in senibus si infirmi sint aliquando .
   cum id ne adulescentes quidem effugere possint ˙

⌐K¬ RESISTENDUM LAELI ET SCIPIO SENECTU ||

Yet many old men are so feeble . that they can perform no
   function that duty or indeed life demands ˙
But that failing is not peculiar to old age . it is a general
   characteristic of ill health ˙
How weak was the son of P. Africanus who adopted you .
   what feeble health he had or rather no health at all ˙
Were it not for this . he would have stood out as the second
   luminary of the state ˙
For to his father's greatness . he had added a more abundant
   instruction of the intellect ˙
What wonder therefore in old men if they are weak some
   times. when even the young cannot escape it ˙

⌐K¬ TO RESIST OLD AGE LAELIUS AND SCIPIO
[is our duty]

Figure 25. Sample from Parkes 1993 showing DISTINCTIO, PUNCTUS INTERROGATIVUS and PARAGRAPHUS.

> SIT UENA TUA BENEDICTA ET LAETARE CUM
> MULI//ERE ADULESCENTIAE TUAE : ~
Sit doctrina tua . et quicumque nascuntur ex ea in benedic-
ti|one ecclesiae . et laetare cum ea cui ab adulescentia id est
| a primo credendi tempore coniunctus es ·
> CERUA CARISSIMA ET GRATISSIMUS HINNU-
> LUS. | UBERA EIUS INEBRIENT TE OMNI TEM-
> PORE | IN AMORE EIUS DELECTARE IUGI-
> TER ℈ : ~

> LET YOUR FOUNTAINS BE BLESSED AND
> REJOICE WITH THE WIFE OF YOUR YOUTH : ~
Let your teaching . and whoever is born from it be in the
blessing of the church . and rejoice with her to whom
from youth that is from the first time of believing you
have been joined : ~
> AS A MOST BELOVED HIND AND A MOST PLEA-
> SANT DOE LET HER BREASTS INEBRIATE
> YOU AT ALL TIMES DELIGHT CONTINUALLY
> IN HER LOVE ℈ : ~

Cerua carissima siue grauissima ut quidam codices |
habent . sancta est ecclesia . quae serpentinam solet odio
habere et conterere doctrinam . gratissimus hinnu|lus
populus eius uirtutum uarietate delectabilis | et eadem
castae fidei semper aemulatione succensus . | Cuius uberi-
bus inebriamur . Cum utriusque testamenti | paginis con-
tra hereticorum fraudes instruimur . | Cuius pace et amore
iugiter delectari magna est | exercendarum occasio uirtu-
tum : ~

The most beloved or as certain manuscripts have vener-
able hind. is holy church. which is accustomed to hate and
crush the teachings of the serpent . the most pleasant doe
is her people pleasing by variety of virtues . and always
kindled by the same emulation of her chaste faith . By
whose breasts we are inebriated . when by the pages of
both testaments we are instructed against the deceptions
of heretics . by whose peace and love to be continually
captivated is a great occasion of exercising virtues : ~

Figure 26. Sample from Parkes 1993 showing COLON WITH COMMA ELEVATA POSITURA.

OF DAVID   79
8 And why should I, lo complaine of thee, which
knowledge my selfe a sinner ⸮ Wherefore, I do not
murmur against thy most iust iudgements, but I ra-
ther desire this of thee, that of thy mercie, forget-
ting al mine iniquities, thou wouldest graunt vnto
me, that I be not a iesting stocke to foolish wicked
men.
9 Finallie, why should not I hold my peace ⸮ why
should I speake one word ⸮ seeing al these things
come vnto me, not by chance, but by thy knowe-
ledge and pleasure, who doubtles art most good,
and most righteous.

Figure 27. Sample from Parkes 1993 showing PUNCTUS PERCONTATIVUS.

Figure 28. Sample from a German dictionary showing TILDE WITH RING.

Figure 29. Sample from Thompson 1912 showing PUNCTUS ELEVATUS and COMMA ELEVATA.

WYCLIFFITE BIBLE.— ABOUT A.D. 1382

(fore þe lord / *and* þe sonys off aaron þe prest | sholen offre þe blood off hit .
shedynge by | envyron of þe auter . þat is byfore þe dore of þe tabernacle /
and þe skyn of þe hoost | drawyn off · þe grete lemys þei sholen kytte in
gobetes . *and* þei sholen ley fuyr in þe | auter · made byfore þe heep of woode ,
and | þe lemes þat ben kut above ordeynynge / | þe heed þat is . *and* al þat
clevyn to þe mawe · þe entrailes *and* þe feet wasche wiþ | water · *and* þe
prest shal brenne hem upon þe auter into al brente sacrifice *and* sweete
smul | to þe lord / þat ʒif of þe beestes is þe offrynge | þe al brent sacrifice
of sheep or of gete · | he shal offre a loomp of o ʒeer . wiþoute wem . *and*
he shal offre at þe syde of þe | auter þat byholdeþ to þe norþe · byfore | þe
lord / þe blood forsoþe of hit þe sonys | of aaron sholen helden upon þe auter ·
by | envyron / *and* þei sholle dyvyden þe lemes | þe heed *and* al þat cleven
to þe mawe · *and* | leye upon þe woode · to þe whiche fuyr is | to be under-
put . þe entrailes forsoþe *and* þe | feet þei sholen whasche wiþ water . *and*
þe | prest shal brenne alle þynges offred upon | þe auter in to brent sacrifise *and*
most sweet | smul to þe lord / ʒif forsoþe off þe briddes | þe offrynge of brent
sacrifice were to þe | lord of turtris or colvyr briddys · þe prest)

Figure 30. Sample from Thompson 1912 showing PUNCTUS ELEVATUS and COMMA ELEVATA.
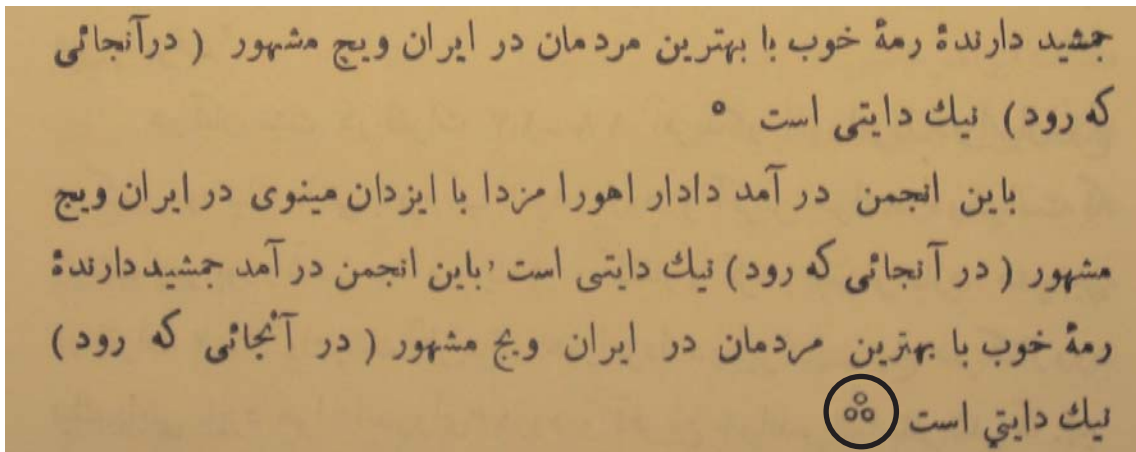
Figure 31. Sample from a Persian Avesta showing LARGE ONE RING OVER TWO RINGS PUNCTUATION.
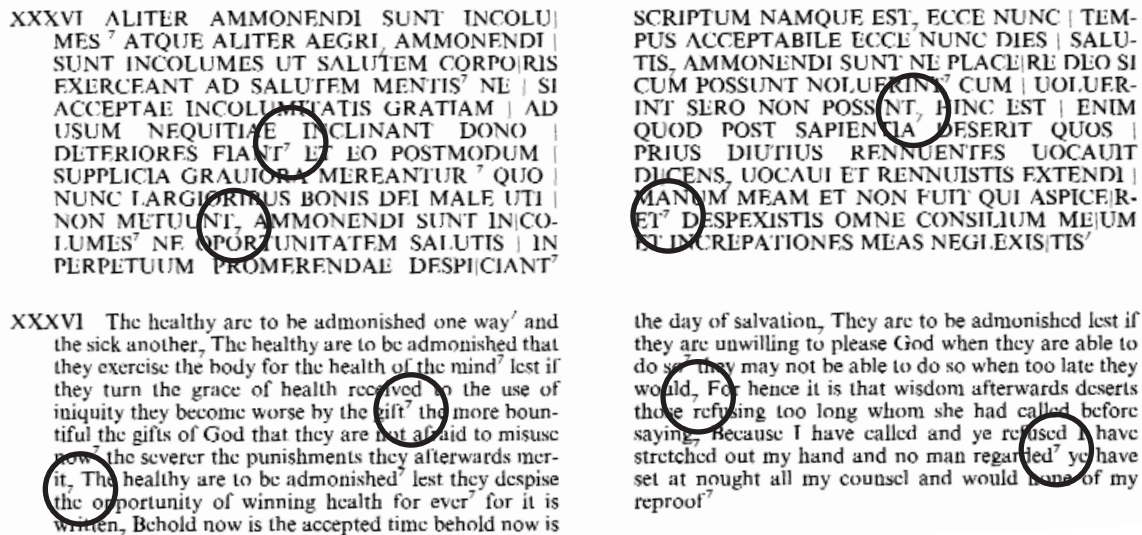
Figure 32. Sample from Parkes 1993, showing POSITURA and SIMPLEX DUCTUS.
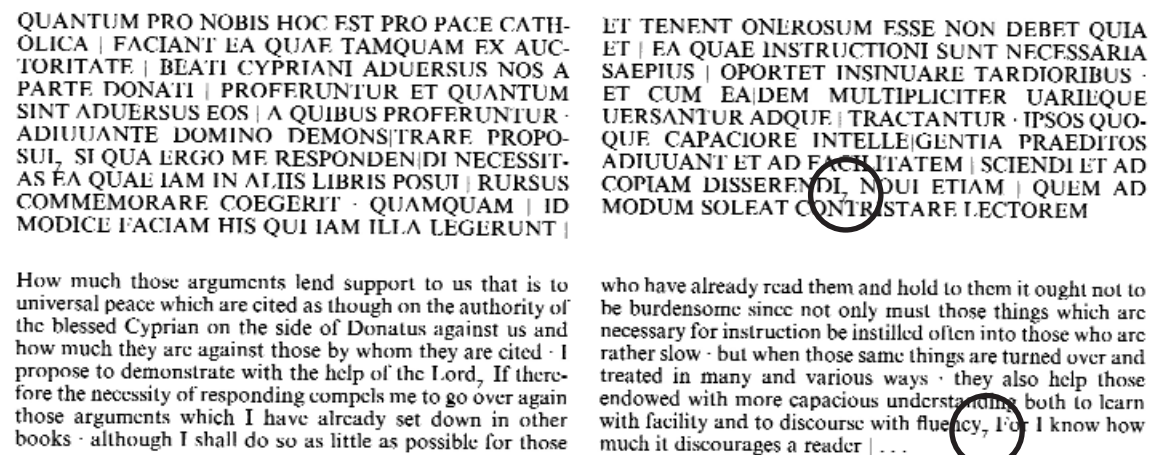
Figure 33. Sample from Parkes 1993, showing POSITURA.

Beneventan

(*a*)　quãre ·,·
　　　quid hoc fecisti ·,·
　　　ũnde estis ·,·

(*b*)　sĩc respondes pontifici ·,·
　　　nũmquid reppulit dominus
　　　　　populum sũum ·,·
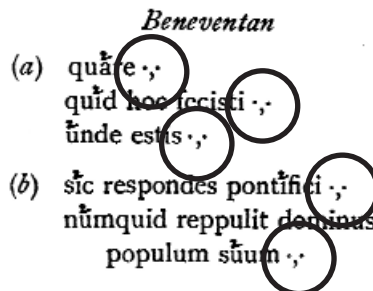
Figure 34. Sample from Parkes 1993, showing TWO DOTS OVER COMMA POSITURA.

By contrast Petrarch used the pointed *virgula* where an interpolated statement has a different relationship to other statements in the immediate context:

Numquam tam iuuenis / numquamque tam gloric cupidus / fui / quod interdum me fuisse non inficior. quin malucrim bonus esse quam doctus .⁹⁵

Never so young / and never so greedy for fame / have I been / I do not deny having been that occasionally. but that I have chosen to be good rather than learned.

Figure 35. Sample from Parkes 1993, showing VIRGULA INTERPOLATIVA.

Il nous faut maintenant compléter l'exposé de Barzizza en signalant ses principales omissions.

J'ai déjà noté l'absence du point d'exclamation (*punctus admirativus*); il était, il est vrai, d'introduction toute récente, puisque le plus ancien exemple actuellement connu figure dans un manuscrit copié peu avant sa mort (1406) par Coluccio Salutati qui en fut peut-être l'inventeur; ce signe se présentait ainsi: !

Figure 36. Sample from Ouy 1987, showing PUNCTUS EXCLAMATIVUS.

signe de renvoi　Any sign used to associate matter in the text with material added in the margin, and especially a passage omitted from the text by the original copyist. ∴ was a sign used by printers.

Figure 37. Sample from Parkes 1993, showing SIGNE DE RENVOI.

Here, used to refer to the base of a 'tick'-shaped stroke. ⁊ ·⁊ ·, :⁓ :⁊ ·,· .;. Originally a critical sign (or *nota*, q.v.), subsequently a punctuation symbol which indicated the end of a section of text; the opposite of a *paragraphus*. For variants see plates 25, 47. :⁓ was employed by insular scribes at the end of a paragraph in a series of paragraphs or texts (e.g. annals) to imply that some continuation was to be expected to complete this series.

Figure 38. Description of various *positurae* from Parkes 1993, showing POSITURA, COLON WITH COMMA ELEVATA POSITURA, COLON WITH POSITURA, TWO DOTS OVER COMMA POSITURA, and THREE DOTS WITH COMMA POSITURA.

Figure 39. Sample from Parkes 1993, showing PUNCTUS EXCLAMATIVUS and PUNCTUS INTERROGATIVUS.

Figure 40. The same sample from Parkes 1993, showing PUNCTUS EXCLAMATIVUS and PUNCTUS INTERROGATIVUS.

THEOLOGICAL TRACTS.- A.D. 821

(capti sunt · et cum illis gentiliter convixerunt · cum | adhuc ad romaniaɴ.
iuvenes venerint · si commu|nionem petierint quid eis observandum sit ·
Si convivio solo gentilium · et ęscis immolaticis usi | sunt · possunt ieiuniis
et manus impositione pur|gari · ut deinceps idolaticis abstinentes · sacramɛ.ʌ
torum *christi* possint esse participes · Si autem | aut idolum adoraverunt · auʈ
homicidiis vel forni|cationibus contaminati sunt · ad communionem)

Figure 41. Sample from Thompson 1912, showing MIDDLE COMMA alongside MIDDLE DOT.

of Seville. An examination of the abbreviations of Monte
Cassino 205 saec. xi suffices to convince us that many of them
are not of the traditional Beneventan stock. The constant use
of symbols like ⊦⊦= *enim,* 7 = *et,* ÷ = *est,* ħ = *hoc,* ·ɪ· = *id est,*
ꝗ = *quod,* ſ· = *sed,* ꝼ = *secundum* and of the form nōe = *nomine*
speaks for Insular influence. By a fortunate coincidence the
Insular origin of the archetype can be attested by other means.

Figure 42. Sample from Loew 1914, showing LONG S followed by MIDDLE COMMA
with the reading *sed* 'but'.

```
10'  n70a,    12'  égi gašan-ki-Zabalam(ZA-MÚŠ-"MÚ[RU")]⌈ki⌋ na-ám-tag-g[a-a-ni
     (+)n70b,  2'      imin a-rá imin    na- (leer)]
     n7la,    11   égi⌋ ⌊gašan-k°i=Zabalam(ZA-M[ÚŠ-"MÚRU"/UNUG)ki                    ]
     BL n205,  6'  [égi gašan]-[ki⌋-Zabalam(⌊ZA⌋-MÚŠ-UNUG)ki  na-[                    ]
```
```
11'  n70a,    13'  égi gašan-Ḫur-sag-[kalam-ma⌐ [n]a-ám-tag-g[a-a-ni imin a-rá
     (+)n70b,  3'      imin    na- (leer)]
```

Figure 43. Sample showing the use in transcription of Babylonian of BOTTOM LEFT HALF SQUARE BRACKET
and BOTTOM RIGHT HALF SQUARE BRACKET.

Here and for the rest of the paper, $\ulcorner \cdot , \cdot \urcorner$ is the Gödel pairing function, *i.e.*, a definable bijection between $\kappa \times \kappa$ and κ. Similarly, $\ulcorner \cdot \urcorner : \kappa^n \to \kappa$ is a definable bijection between κ^n and κ.

We define a colouring of $\mathbb{C}_\kappa^\rho$ as follows:

$$c_{A,\vec{\mu}}(f) := \begin{cases} 1 & \text{if } \ulcorner [f_0]_{\mu_0}, \ldots, [f_{n-1}]_{\mu_{n-1}} \urcorner \in A, \\ 0 & \text{otherwise.} \end{cases}$$

For $f : \kappa \to \lambda$ define $\hat{f} : \kappa \to \lambda^\kappa / \nu$ by $\hat{f}(\ulcorner \vec{\alpha} \urcorner) := [\ulcorner \vec{\beta} \urcorner \mapsto f(\ulcorner \vec{\alpha} ^\frown \vec{\beta} \urcorner)]_\nu.$

$$f(\ulcorner \overrightarrow{[x]}_{\mathbf{germ}_\mu} ^\frown \overrightarrow{[y]}_{\mathbf{germ}_\nu} \urcorner) \le g(\ulcorner \overrightarrow{[x]}_{\mathbf{germ}_\mu} ^\frown \overrightarrow{[y]}_{\mathbf{germ}_\nu} \urcorner).$$

$$[\ulcorner \vec{\beta} \urcorner \mapsto f(\ulcorner \overrightarrow{[x]}_{\mathbf{germ}_\mu} ^\frown \vec{\beta} \urcorner)]_\nu \le [\ulcorner \vec{\beta} \urcorner \mapsto g(\ulcorner \overrightarrow{[x]}_{\mathbf{germ}_\mu} ^\frown \vec{\beta} \urcorner)]_\nu,$$

which by definition of $f \mapsto \hat{f}$ translates to $[\hat{f}]_\mu \le [\hat{g}]_\mu.$

Figure 44. Sample of mathematical text showing the showing SQUARE BRACKETS and QUINE CORNERS

In what follows, we frequently conserve notation by writing $V_i(x,r)$ when we should more properly write $V_i(\lfloor x \rfloor, r)$.

The smallest integer δ_u satisfying the inequality of the statement is $\delta_u = \left\lceil \frac{2h_u - h_i - 2}{3} \right\rceil$, where as usual $\lceil \alpha \rceil$ indicates the least integer greater than or equal to α and $\lfloor \alpha \rfloor$ indicates the largest integer less than or equal to α. Now show these in a display:

$$\lceil \alpha \rceil, \quad \lfloor \alpha \rfloor, \quad \delta_u = \left\lceil \frac{2h_u - h_i - 2}{3} \right\rceil$$

$$(4.1) \qquad\qquad T = \text{conv} \left(\bigcup_{i=1}^{\lfloor \log_2 n \rfloor} \frac{1}{i} K_{2^i} \right).$$

Let $s = \lfloor \log_2 n \rfloor$ and $m = \lfloor \log_2(\sqrt{n}) \rfloor \simeq s/2.$

Figure 45. Sample of mathematical text showing the showing FLOORS and CEILINGS
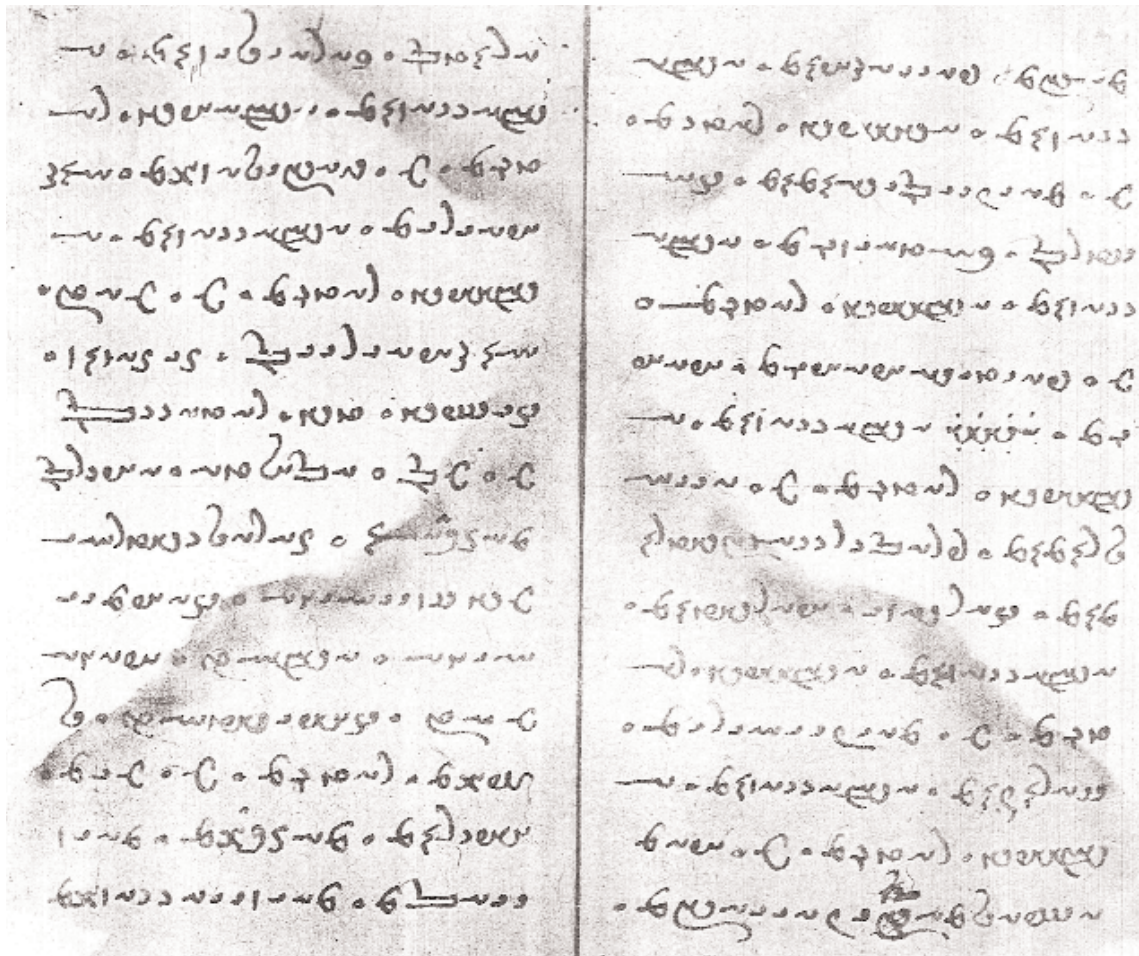
Figure 46. Sample of an Avestan Yasna manuscript Manuscript K7, fol 29v showing the use of the RING POINT. Taken from **http://avesta.org/gifs/samples.htm**.

A. Administrative
1. Title
Proposal to encode Medievalist and Iranianist punctuation characters to the UCS
2. Requester's name
Michael Everson, Peter Baker, Marcus Dohnicht, António Emiliano, Odd Einar Haugen, Susana Pedro, David J. Perry, Roozbeh Pournader
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2007-01-09
5. Requester's reference (if applicable)
6. Choose one of the following:
6a. This is a complete proposal
Yes.
6b. More information will be provided later
No.

B. Technical – General
1. Choose one of the following:
1a. This proposal is for a new script (set of characters)
No.
1b. Proposed name of script
1c. The proposal is for addition of character(s) to an existing block
Yes
1d. Name of the existing block
Supplementary Punctuation and a new Avestan block
2. Number of characters in proposal
48 (40, 8).
3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)
Category B.1.
4a. Is a repertoire including character names provided?
Yes.
4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?
Yes.
4c. Are the character shapes attached in a legible form suitable for review?
Yes.
5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson.
5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson, Fontographer.
6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes.
6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.
7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?
Yes.
8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.
See above.

C. Technical – Justification
1. Has this proposal for addition of character(s) been submitted before? If YES, explain.
Yes, some of the characters have. See N3178.
2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes.

2b. If YES, with whom?

The authors are members of the user community.

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Medievalists, Latinists, Slavicists, Iranianists, and other scholars.

4a. The context of use for the proposed characters (type of use; common or rare)

Used historically and in modern editions.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

Scholarly publications.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Accordance with the Roadmap. Keep with other punctuation characters.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

Yes.

9b. If YES, is a rationale for its inclusion provided?

Yes.

9c. If YES, reference

It is possible to analyse the four German punctuation characters as being decomposable, but diacritics are not normally applied to punctuation and these characters are best treated as undecomposable units.

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

Yes.

10b. If YES, is a rationale for its inclusion provided?

Yes.

10c. If YES, reference

See §3.6 above. Also the PUNCTUS PERCONTATIVUS looks like ARABIC QUESTION MARK but the first has a character property ON and the second AL.

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?