Format of a PIL Running Text File

H. J. Bosman C. J. Sikkel

Version 2.5 of January 15, 2015

Abstract

This document describes the so-called 'running text' files of the Peshitta Institute Leiden for the CALAP and Turgama projects. It defines the format of the source files with the Aramaic text and serves as a standard for programmers and maintainers of the text. In case of uncertainties in the descriptions in Section 2, the formal description in Section A of the appendix will be decisive. This report is part of the documentation of both CALAP and Turgama.

Contents

1 Introduction

1.1 The Need for a Standard

For the automatic processing of a computer text, a clear definition of the technical format of that text is essential. It is also important that this definition is written down explicitly in advance:

1. A computer program needs to know exactly what signs it can expect in the text and in what order. Unlike a human reader, it does not have the flexibility and understanding to make an 'educated guess' when encountering something unexpected.

2. Therefore, the developers of the software also need to know exactly what they can expect in the text, so they can design the software accordingly, without having to consult the other parties involved.

3. The maintainers of the text files also need to have such a written standard, so that they know what format the software developers expect, without need for consultation.

1.2 The PIL Format

At the start of CALAP, the 'running text' files already exhibited a consistent format for chapter headings, chapter endings, verse numbers, verse endings, Syriac characters and textual variations.[1] Up to now, this has been partly a matter of convention, as the format has not been written down formally. The present paper is an attempt to do exactly that. It follows the format of the original CALAP files[2] and thus presents a description of the *status quo* rather than something new. A few deviations from this *status quo* will be proposed, however.

First of all, the use of spaces and ends of lines (newlines) is defined in a flexible way, allowing for a transparent layout. Moreover, the possibility is added of including comments anywhere in the running text. This is an extension, rather than a change, of the format. The CALAP files as they have been up to now are entirely compatible with this extension.

[1] A description of the format has been given in [BM89].

[2] This format has been gathered from one file: the July 1999 version of `1-2k-run.txt`, the PIL electronic text of the book of Kings.

Secondly, the format of variants has been defined more strictly and consistently. This stricter definition will affect only a small number of cases in the running texts. A real change is the use of the ampersand (`&`) for the closing angular bracket (`>`) to indicate 'et alii manuscripti'. This was done to avoid ambiguities with comments, which are denoted by angular brackets.

Finally, the encoding scheme for diacritics and punctuation marks has been expanded, allowing the distinction of a wider variety of signs than the running text files in their original form exhibit. A real change here is the token of the *seyame*, which is no longer the colon (`:`), but the double quote (`"`). This is done because of mnemotechnical reasons and in order to attain correspondence with the *rahta*.

At the start of the Turgama project in 2005, the decision was made to extend the format so that it could be used for both Peshitta and Targum texts. This was accomplished by adding codes for the graphemes *ḥolem* and *sin*. The code chosen for the *shewa* was already part of the standard, because it features in some of the composite punctuation marks. Henceforth we shall use the term 'Aramaic' in an inclusive sense covering both (Babylonian) Aramaic and (Peshitta) Syriac.

2 Description

The following description is hierarchically structured in a top-down fashion. A text file is described as consisting of chapters, a chapter as consisting of verses, and so on. On any given level of the description, cross references to relevant descriptions on other levels are given.

2.1 Text File

1. A *text file* contains the chapters (Section 2.2) of one or more books.
2. The transition from one book to the next is not indicated by any special marker, but follows from the change in the contents of the chapter headings (Section 2.3).
3. Although it is common practice in the CALAP files to separate chapters by blank lines, the present format no longer requires this.

2.2 Chapter

A *chapter* consists of:

1. A chapter heading (Section 2.3);
2. One or more verses (Section 2.4).

In the format used thus far, no newlines were allowed between the verses of a chapter. In the present proposal, placing each verse on a separate line is not only permitted,[3] but actually encouraged, because it greatly facilitates reading and editing (see Figure 1 on the following page).

[3] Albeit not required: programs must accept chapters with 'run-on' verses as in the old format.

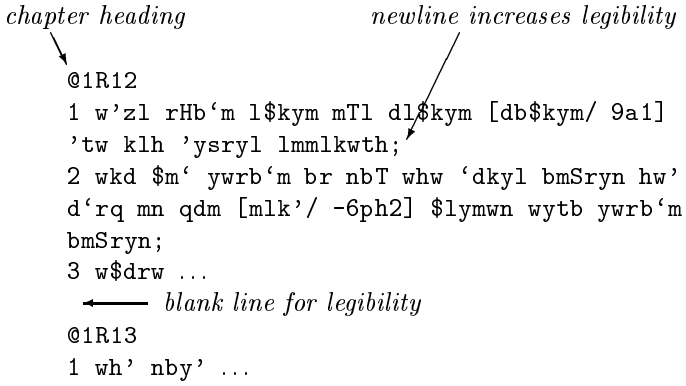

Figure 1: Example of the format of a chapter

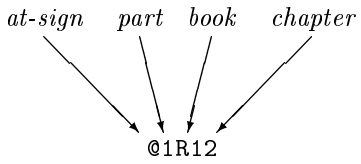

Figure 2: Example of the format of a chapter heading

2.3 Chapter Heading

A *chapter heading* consists of:

1. An *at sign* (`@`);
2. A number indicating the part of the book (if applicable);
3. A string of alphabetic characters identifying the book;
4. A number identifying the chapter.

There are no spaces between these elements (see Figure 2).

2.4 Verse

A *verse* consists of:

1. The verse number (digits only), followed by a space (Section 2.10);
2. The annotated running text of that verse (Section 2.5), which can be empty;
3. A *semicolon* (`;`), which is followed by a space if another verse follows on the same line.

An example is given in Figure 3 on the following page.

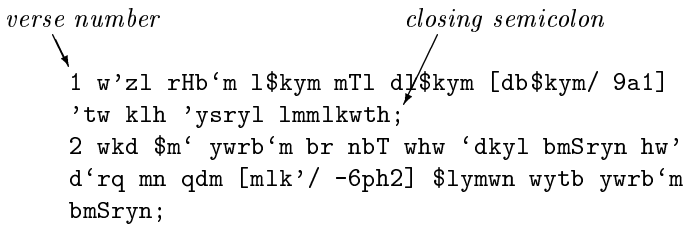

Figure 3: Example of the format of a verse (1R 12:1–2)

2.5 Annotated Running Text

Annotated *running text* consists of zero or more of the following elements:

1. Aramaic writing (Section 2.6).
2. Textual variant (Section 2.7).
3. Manuscript boundary (Section 2.8).
4. Comment, concerning corrections et cetera (Section 2.12).

2.6 Aramaic Writing

A stretch of *Aramaic writing* consists of one or more of the following characters

```
" # $ ' * . : @ A E F H O S T ^ _ ‘ a b d e g h i k l m n o p q r s t
u w y z ~
```

or digraphs

```
#! #" #, #. #/ #: #\ #_ =! =" =, =. =/ =: =\ =_ ^! ^" ^, ^. ^/
^: ^\ ^_
```

Note that spaces are not considered part of the writing proper; they only serve to separate stretches of writing. In practice, a single stretch of Aramaic writing can be:

1. An Aramaic word, including any diacritics or punctuation marks which may be attached to it;
2. Punctuation which is *not* attached to an Aramaic word, but is separated from it by white space (Section 2.10).

Section B of the appendix shows the transliteration used in the notation of Aramaic writing. This transliteration consists entirely of characters from the set mentioned above.

2.7 Variant

The definition of a *variant* has been derived from the format that was *de facto* used in the electronic texts. This format appears to be based on—but is slightly different from—the description in [BCJW97, xvi–xvii]. Our definition is as follows.

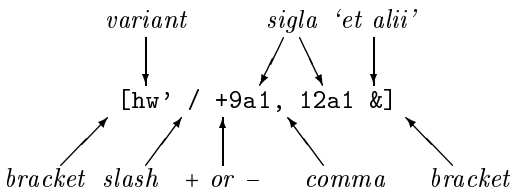

Figure 4: Example of the format of a variant (2R 1:13)

1. A variant is enclosed within square brackets: `[ ... ]`.
2. A slash (/) separates the variant reading proper from the list of manuscripts pertaining to the reading.
3. The variant reading consists of one or more stretches of Aramaic writing (Section 2.6). The last one of these may end with a dash (-), to indicate a prefix.
4. The list of manuscripts may start with one plus or minus. A plus marks a variant that the following manuscripts have more than the textus receptus; a minus indicates a variant that the manuscripts have less than the textus receptus.
5. The manuscripts are indicated by sigla (Section 2.9). The sigla must be separated by comma's.
6. The last siglum in the list may be followed by an *ampersand* (`&`), indicating 'et alii manuscripti'.

The notation of variants is discussed at greater length in Section 3. An elaborated example of a variant is given in Figure 4.

2.8 Manuscript Boundary

2.8.1 Notation

The possibility of marking the boundaries of a manuscript is a later addition to the format. The boundaries are marked in the following manner:

1. The statement is enclosed within double square brackets: `[[ ... ]]`.
2. The statement consists of a plus or a minus followed by a list of manuscript sigla. A minus indicates that the named manuscripts cease to be a witness from this point on in the text. A plus reintroduces the named manuscripts as witnesses after they had been excluded earlier in the text, for instance at the end of a lacuna.
3. The manuscripts are indicated by sigla (Section 2.9), which are separated by comma's.

As an example, to indicate that $6ph2$ and $8h4$ are lacking from this point on, one would write `[[-6ph2,8h4]]`. At the start of the text file, all manuscripts are assumed to be present. Thus the semantics of files without manuscript boundaries is preserved.

2.8.2 Lacunae

Manuscript boundary indicators can be used to delimit a lacuna in a particular manuscript. As the manuscript is assumed to be present from the start of the text, it suffices to mark the place where the manuscript ends and the place where it resumes. The lacuna in $8h4$ spanning 1R 22:20–34, for instance, can be represented as follows.

```
20 w'mr [[-8h4]] mry' mnw n$dlywhy l'Hb wnsq wnpl brmt gl‘d hn'
   'mr hkn' whn' 'mr hkn';

   <verses 21 to 33 follow>

34 wgbr' Hd [[+8h4]] $d' hw' bq$t' lqwblh tmym'yt wmHyhy lmlk'
   d'ysryl byt dbq' d$ryn' w'mr lmrkbnh 'hpk 'ydk w'pqyny mn
   m$ryt' mTl dmTywny H"bl' dmwt';
```

2.8.3 Fragments

A fragment is indicated in a similar, but inverse manner. As the manuscript is assumed to be present from the start of the text, we need to state its absence at the beginning of the file. The fragment is then demarcated by the place where it starts and the place where it ends. The fragment of $9k4$ containing 2R 10:14–25, for instance, can be represented as follows.

```
 0 [[-9k4]]; <top of the file>

   <text of 1R 1:1 to 2R 10:13>

14 w'mr=.  'Hwdw 'nwn kd H"yyn w'Hdw 'nwn wnksw 'nwn [[+9k4]]
   w$dw 'nwn bgwb' 'r"b‘yn wtr"yn gbr"yn wl' $bq mnhwn 'n$;

   <verses 15 to 24 follow>

25 wkd gmrw lm‘bd dbH"' w‘"lwt' 'mr yhw lr"hT' wlgnbr"'
   [[-9k4]] ‘wlw Hrwbw 'nwn w'n$ l' npwq mnhwn wHrbw 'nwn bpwm'
   dsyp' w'rmyw 'nwn r"hT' wgnbr"' w'zlw ‘dm' lqryt' dbyt b‘l';
```

2.9 Manuscript Siglum

A siglum referring to a manuscript consists of:

1. An optional prefix in the form of a number indicating the century of the addition and a slash;
2. A number (the century of the manuscript);
3. One or more alphabetic characters (the type of the manuscript);
4. A number (the ordinal number of the manuscript)

`N`	Nestorian text tradition: one capital letter only
`9a1`	common siglum: number-letter(s)-number
`10c1*`	siglum with subscript `*`
`8/5b1`	eighth century addition to `5b1`
`6h18<vid>`	siglum followed by the comment `<vid>`

Figure 5: Examples of the format of a manuscript siglum

5. A subscript (optional). There may be only one subscript in a siglum. The following subscripts are accepted: `* c fam mg txt`.

6. Alternatively, a siglum may consist of a single capital letter (presently only N) to refer to an entire manuscript tradition.

Anything pertaining to sigla, which does not fit into the above schema, such as indication of uncertainties, is to be included as a comment (Section 2.12) following the siglum (see Figure 5).

2.10 Space

1. A *space* may be written as a single space character, a tab character or a newline (Section 2.11), but also as any number of subsequent space characters, newlines and tab characters. Spaces, tabs and newlines may be used freely in order to produce an easily readable layout.

2. A comment (Section 2.12) also counts as a space, and therefore has the same effect: it separates tokens and words.

3. There must be a space *between* Aramaic words (Section 2.6), whereas between other tokens (e.g., inside a variant) they may be omitted.

4. There may be no space *within* an Aramaic word, within a chapter heading or within a manuscript siglum.

2.11 Newline

This standard treats the end of a line (newline) in exactly the same way as a space.[4] Newlines may occur anywhere in the running text,[5] as long as they do not appear within an Aramaic word, chapter heading or manuscript siglum. Thus, it is permitted to use newlines within and around comments and after every verse. Since it greatly improves readability, the latter practice is actually encouraged.

[4] A technical note: there is a difference between UNIX and DOS in the way a newline is marked. Both notations are accepted by this format. For practical reasons, the project files are now stored following the UNIX convention.

[5] This constitutes an extension of the format used thus far, in which a newline could only be used to end a chapter or a chapter heading. In the present format, the newline has no special meaning and is treated as any white space character.

P	the text of the prefix
S	the set of manuscripts to which the variant pertains
V	the text which is the subject of the variant
W	the text left unaltered by the variant

Table 1: Symbols

2.12 Comment

1. A *comment* is enclosed within angular brackets: < ... >.
2. There may therefore be no angular brackets *within* a comment.
3. Comments are treated as if they were spaces. They may occur anywhere in the file where a space might also appear (Section 2.10).
4. There may therefore be no comments *within* an Aramaic word, within a chapter heading or within a manuscript siglum.
5. Comments may contain newlines (Section 2.11); they may also be preceded or followed by newlines.

3 Variants

3.1 Introduction

It is fairly straightforward how the format defined in this document serves the encoding of Aramaic texts. The recording of variants, however, requires further explanation. In this section we shall describe more precisely what is meant by the notation defined in Section 2.7.

We distinguish two types of variants: word variants and prefix variants. The word variant is the general case, which arises when two texts have different words. A word, in this context, is a stretch of Aramaic writing as defined in Section 2.6. The prefix variant is a shorthand notation for the word variant in which two single words differ in their prefix only.

In what follows, we use symbols to explain the notation of variants. The meaning of these symbols is summarised in Table 1. In a prefix variant, W stands for the part that remains when the word is stripped of its prefix. The prefix is denoted by P, which is confined to a single word.

3.2 Addition

The notation $W[V/+S]$ indicates an addition to the main text. The main text reads W, and the manuscripts in S read $W\,V$. Note the space between W and V. In case of a prefix variant, the notation is $W[P\text{-}/+S]$. Here the main text reads W, and the manuscripts in S read PW. Note that PW is one word. An example of an addition is found in 1R 1:5, where $9a1$ has tacked on a form of *to be*.

```
5 w'dwny' br Hgyt mtrwrb [hw'/ +9a1] w'mr 'n' 'mlk w‘bd lh
  mr"kbt' wpr"$' wHm$yn gbr"yn drhTyn hww qdmwhy;
```

A prefix addition is found in 1R 2:27, where a *dalath* is added in $6h18$ and $11c1$.

```
27 w'pqh $lymwn l'bytr dl' nhw' khn' lmry' dn$tml' ptgmh dmry'
   dmll 'l byt [d-/ +6h18, 11c1] 'ly b$ylw;
```

We shall see below that a prefix addition is a shorthand notation for the following type of substitution.

$$W[P\text{-}/+S] \equiv W[PW/S] \tag{1}$$

When W has a length of more than two characters, the notation of a variant as a prefix addition is conciser than a substitution.

3.3 Deletion

The notation $W[V/-S]$ indicates a deletion from the main text. The main text reads $W\,V$, and the manuscripts in S read W. Note the space between W and V. In case of a prefix variant, the notation is $W[P\text{-}/-S]$. Here the main text reads PW, and the manuscripts in S read W. Note that PW is one word. An example of a deletion is found in 1R 1:4, where $9a1$ has two words left out.

```
 4 w'lymt' $pyr' [hwt bHzwh/ -9a1] Tb whwt lmlk' m$m$nyt'
   wm$m$' lh wmlk' l' yd'h;
```

A prefix deletion is found in 1R 9:20, where a *waw* is missing in $7a1$.

```
20 klh [w-/ -7a1] 'm' d'$tHrw mn 'mwr"y' wH"ty' wpr"zy' wH"wy'
   wyb"wsy' dl' hww mn bn"y 'ysryl;
```

We shall see below that a prefix deletion is a shorthand notation for the following type of substitution.

$$W[P\text{-}/-S] \equiv PW[W/S] \tag{2}$$

When W has a length of more than two characters, the notation of a variant as a prefix deletion is conciser than a substitution.

3.4 Substitution

The notation $V_1[V_2/S]$ indicates a substitution. The number of words in V_1 must equal the number of words in V_2. The substitution is characterised by the absence of an edit operator (+ or –) after the slash. The meaning of a substitution is by definition that of a deletion immediately followed by an addition.

$$V_1[V_2/S] \stackrel{\text{def}}{=} [V_1/-S][V_2/+S] \tag{3}$$

Again, with the restriction that the number of words in V_1 must equal the number of words in V_2. Only in cases where the number of words in the verse before the substitution is less than the number of words in V_2, can violation of this rule be detected automatically.

In case of a prefix variant, the notation is $P_1W[P_2\text{-}/S]$. Here the condition must be met that the length of P_1 in characters equals the length of P_2. The

Notation	Text	S
$W[V/-S]$	$W\,V$	W
$W[V/+S]$	W	$W\,V$
$V_1[V_2/S]$	V_1	V_2
$W[P\text{-}/-S]$	PW	W
$W[P\text{-}/+S]$	W	PW
$P_1W[P_2\text{-}/S]$	P_1W	P_2W

Table 2: Notation of Variants

meaning of this type of substitution is defined in terms of the corresponding full substitution.

$$P_1W[P_2\text{-}/S] \stackrel{\text{def}}{=} P_1W[P_2W/S] \tag{4}$$

Again, with the restriction that the length of P_1 must equal the length of P_2. Only in cases where the length of P_1W is not greater than the length of P_2, can violation of this rule be detected automatically. Unless W is exactly one character long, the notation of a variant as a prefix substitution is conciser than a full substitution.

An example of a substitution is found in 1R 6:12, where $7a1$ has a different word order.

```
12 byt' hn' dbnyt 'n thlk bqy"my wdy"ny tTr wt‘bd [t‘bd wtTr/
   7a1] klhwn pwq"dny wthlk bhwn 'qym pt"gmy ‘mk d'mrt ldwyd
   'bwk;
```

A prefix substitution is found in 1R 7:7, where $7h10$ has a *lamad* in stead of a *dalath*.

```
 7 w'sTw' dkwrsy' [l-/ 6ph2, 7h10] dd'n hw' tmn 'sTw' ldyn' ‘bd
   wqrmh b'r"z' mn $t'"swhy w‘dm' l$m"why;
```

3.5 Evaluation Order

The order of evaluation of the variants is from left to right. This means that in case of a substitution the variant is applied to the text which results from the evaluation of all variants left to this substitution. This opens the door to contructions in which a textual variation is given as a variant of a variant. It is questionable, however, whether it is prudent to make use of this possibility, as it makes the text less legible and more susceptible to coding errors. This is the more so, as the number of manuscripts involved increases. This concludes our discussion of the variants. An overview of the notation is given in Table 2.

A Formal Description

A formal specification of the file format is given with the grammar below. The notation used has been derived from the Backus-Naur Form and uses the following additional metasymbols.

`term +`	one or more instances of `term`
`term ?`	zero or one instance of `term`
`{ ... }`	zero or more repetitions of what is enclosed
`[ ... ]`	regular expression matching a single character

```
text            =  chapter +
chapter         =  heading verse +
heading         =  '@' number ? string number
verse           =  number word + ';'
word            =  writing | variant | boundary
variant         =  '[' writing + '-' ? '/'
                   operation ? siglum { ',' siglum } '&' ? ']'
writing         =  grapheme +
boundary        =  '[[' operation siglum { ',' siglum } ']]'
grapheme        =  letter | word_bound |
                   vowel | composite | pericope
letter          =  ['bgdhwzHTyklmns‘pSqrF$t]
word_bound      =  ["#^~]
vowel           =  [:AEOaeiu]
composite       =  [#=^][!",./:\_]
pericope        =  [*.@_o]
operation       =  '+' | '-'
siglum          =  tradition | ms
tradition       =  [[:upper:]]
ms              =  number string number subscript ?
string          =  [[:alpha:]] +
number          =  [[:digit:]] +
subscript       =  '*' | 'c' | 'fam' | 'mg' | 'txt'
```

B Grapheme List

This section lists the graphemes that occur in Aramaic writing with their encoding in ASCII characters. Where possible, the corresponding Unicode character code for Syriac is also listed. For the differentiation of the graphemes into letters, diacritics and punctuation marks, we refer to the WIVU manual page `alphabet(5)`.

B.1 Letters

’	alaf	U+05D0	U+0710	m	mim	U+05DE	U+0721
b	beth	U+05D1	U+0712	n	nun	U+05E0	U+0722
g	gamal	U+05D2	U+0713	s	semkath	U+05E1	U+0723
d	dalath	U+05D3	U+0715	‘	‘e	U+05E2	U+0725
h	he	U+05D4	U+0717	p	pe	U+05E4	U+0726
w	waw	U+05D5	U+0718	S	tsade	U+05E6	U+0728
z	zain	U+05D6	U+0719	q	qof	U+05E7	U+0729
H	heth	U+05D7	U+071A	r	resh	U+05E8	U+072A
T	thet	U+05D8	U+071B	F	sin	U+05E9 U+05C2	N/A
y	yod	U+05D9	U+071D	$	shin	U+05E9 U+05C1	U+072B
k	kaf	U+05DB	U+071F	t	taw	U+05EA	U+072C
l	lamad	U+05DC	U+0720				

B.2 Diacritics

We distinguish the word-bound diacritics, like the ‘diacritical dot above’, from the letter-bound diacritics, like the *qushshaya*.[6]

B.2.1 Word-bound Diacritics

"	seyame	U+0308
#	diacritical dot below	U+0323
^	diacritical dot above	U+0307
~	abbreviation mark	U+070F

The Syriac abbreviation mark is a line above the last few letters of a word. For example, `t~$bw` for `t$bwHt’`[Kir01, §1.11]. The `~` is put before the first letter with the line over it.

B.2.2 Letter-bound Diacritics

The non-vocalic diacritics are encoded using a combination of `#`, `=`, or `^` (indicating whether the diacritic is below, level with, or above the letter) and a character reminiscent of the form of the diacritic. From Table 3 on the next page, we have the following non-vocalic letter-bound diacritics. The philological name of some of the diacritics is unclear, even in Segal’s book on the diacritics [Seg53].

#,	rukkakha	U+0742
#"	unclear	U+0324
#!	unclear	U+0744
#_	linea occultans infera	U+0331
^,	qushshaya	U+0741
^!	unclear	U+0743
^_	linea occultans supera	U+0304

[6] In earlier versions of this document, no distinction was made between the *qushshaya* and the ‘diacritical dot above’, and likewise between the *rukkakha* and the ‘diacritical dot below’. In many manuscripts, however, the two kinds of dot are distinguishable, because the *qushshaya* and *rukkakha* are rendered with distinctly smaller dots (observation made by Wido). They are therefore to be considered as graphemes different from the diacritical dots.

[7] The philological name of some of the diacritics is unclear, as noted in Section B.2.2

Form	#	=	^	Char
◌̇	rukkakha U+0742		qushshaya U+0741	’
◌̈	unclear[7] U+0324			"
◌̣	unclear[7] U+0744		unclear[7] U+0743	!
◌̱	linea occultans infera U+0331		linea occultans supera U+0304	-

Table 3: Non-vocalic letter-bound diacritics

The vocalic diacritics are represented by a single character.

:	shewa	U+05B0	N/A
A	qamets, zeqapa, zeqofo	U+05B8	U+0733–U+0735
E	tsere, revasa karya	U+05B5	U+0739
O	ḥolem, rewaha	U+05B9	U+073F
a	pataḥ, petaḥa, petoḥo	U+05B7	U+0730–U+0732
e	segol, revasa arrika, revoso	U+05B6	U+0736–U+0738
i	ḥireq, hevoso	U+05B4	U+073A–U+073B
u	qubbuts, esoso	U+05BB	U+073D–U+073E

The *hevasa* U+071D U+073C, the *esasa allisa* U+0718 U+073C and the *esasa rewiha* U+0718 U+073F are rendered as y#, w# and w^ respectively.

B.3 Punctuation Marks

We distinguish the punctuation marks proper, which are encoded with a character combination, from the pericope markers, which are represented by a single character.

B.3.1 Interpunction

The punctuation marks are encoded using a combination of #, =, or ^ (indicating whether the diacritic is below, level with, or above the letter) and a character reminiscent of the form of the mark. From Table 4 on the following page, we can list the following punctuation marks.

Form	#	=	^	Char
· 	menachta U+0702	pasuqa U+002E	paquda U+0701	.
‥ 			rahta U+0705	"
⁛ 		elaya U+0707		/
׃ 	metkashpana U+0704	shewaya U+003A	taksa U+0703	:
∴ 	tahtaya U+0709	unclear U+0706	unclear U+0708	\

Table 4: Punctuation marks proper

#.	menachta, meshalyana (ES), metdamrana, samka	U+0702
#:	metkashpana (ES)	U+0704
#\	tahtaya, metkashpana (WS), meshalyana (WS)	U+0709
=.	pasuqa	U+002E
=/	elaya	U+0707
=:	shewaya (WS), zauga (ES)	U+003A
=\	unclear	U+0706
^.	paquda, metkashpana (ES), meshalyana (ES), etsyana, meshalana?	U+0701
^"	rahta	U+0705
^:	taksa (WS), zauga elaya (ES)	U+0703
^\	unclear	U+0708

B.3.2 Pericope Markers

*	rosette	U+0700
.	common dot in caesuras	U+00B7
@	vignette	U+2722
_	dash in caesuras	U+2014
o	large dot in caesuras	U+2022

References

[Ass89] Association Internationale Bible et Informatique. *Actes du Second Colloque International "Bible et Informatique: méthodes, outils, résultats," Jérusalem, 9–13 juin 1988*, number 43 in Travaux de linguistique quantitative, Paris-Genève, 1989. Champion-Slatkine.

[BCJW97] P. G. Borbone, J. Cook, K. D. Jenner, and D. M. Walter, editors. *Concordance. Pentateuch*, volume V/1 of *The Old Testament in Syriac according to the Peshiṭta version*. Brill, Leiden, 1997.

[BM89] Pier Giorgio Borbone and Francesco Mandracci. An other way to analyze syriac texts. a simple powerful tool to draw up syriac computer aided concordances. In *Actes du Second Colloque International "Bible et Informatique: méthodes, outils, résultats," Jérusalem, 9–13 juin 1988* [Ass89], pages 135–145.

[Kir01] George A. Kiraz. *Meltho: Syriac OpenType Fonts for Windows XP/2000^{TM} and Windows 95/98/ME^{TM}*. Beth Mardutho: The Syriac Institute, Piscataway, NJ, 2001.

[Seg53] J. B. Segal. *The Diacritical Point and the Accents in Syriac*. Number 2 in London Oriental Series. Cumberlege, London, 1953.