

A Reference Guide to the Westminster Leningrad Codex

The J. Alan Groves Center for Advanced Biblical Research
(known as “The Groves Center” for short)
2960 Church Road
Glenside, Pennsylvania 19038-2000
U.S.A.

Document Release 4.20
Monday 25 January 2016

NOTE: The Groves Center's President and Senior Research Fellow, Kirk E. Lowery, Ph.D., has retired from the Groves Center as of November 1st, 2015. . The board of the Groves Center publicly thanks Dr. Lowery for his 17 years of service to the Groves Center (the first half of which were under its previous name, “The Westminster Hebrew Institute”).

Please direct all correspondence to Mr. Stephen K. Salisbury, the Groves Center's Executive Director and Research Fellow. Since our last release, we have changed our e-mail server infrastructure, and some earlier e-mail addresses (specifically “ssalisbury@whi.wts.edu” and “ssalisbury@GrovesCenter.org”) are no longer supported.

Please use the e-mail address Steve@GrovesCenter.org to contact the G.C.

See the **Release Notes** file for specific changes not covered in this document.

General Introduction

The Westminster Leningrad Codex (WLC) is a simple ASCII file, containing roughly twenty-three thousand two hundred lines of text, totaling about three and a half million characters, representing the text of the Hebrew Bible.

The Groves Center is the owner, maintainer, and distributor of the WLC. The editor of the WLC is Mr. Stephen K. Salisbury, the Groves Center's Executive Director and Research Fellow.

In WLC 4.6 and earlier, the file consisted of 23,213 lines of transliterated Hebrew text, one line per verse in the Hebrew Bible. (Note that the verse numbering in the Masoretic Text of the Hebrew Bible does not always line up with the verse numbering in English Bibles.)

Starting with the WLC 4.8, a dozen or so lines of comment text (each line beginning with the character “#”) were added at the beginning of the file. These lines identify the file version, the owner of the file, copyright information, contact name(s) and e-mail address(es), and the date and time of and internal version number of the files used to create the release.

WLC Line Format

Each line of the WLC (after the initial header of a dozen or so lines beginning with “#”) consists of a two-letter book abbreviation, a chapter number, a colon character (“:”), the verse number, a space and then the transliterated Hebrew and/or Aramaic text of that verse. Hebrew and Aramaic text are not distinguished in our transliteration because they are not distinguished in the original manuscripts of the Hebrew Bible itself (or, for that matter, in the modern printed editions).

The Aramaic sections of the Hebrew Bible are:

Genesis 31:47 (4th word)
Jeremiah 10:11
Daniel 2:4 (5th word) – 7:28
Ezra 4:8 – 6:18 and 7:12 – 26

Chapter and verse numbers vary from one to three digits and are displayed without leading zeroes.

Note that the book order is not English (Vulgate), nor Rabbinic (*Miqraot Gedolot*), but the order of the Leningrad Codex. The difference between the Rabbinic order and the Leningrad Codex is slight: a shift of Chronicles from the end of the Hebrew Bible to right before Psalms.

The abbreviations for the names of the books of the Hebrew Bible are:

gn	Genesis	2k	2 Kings	na	Nahum	pr	Proverbs
ex	Exodus	is	Isaiah	hb	Habakkuk	ru	Ruth
lv	Leviticus	je	Jeremiah	zp	Zephaniah	ca	Canticles
nu	Numbers	ek	Ezekiel	hg	Haggai	ec	Ecclesiastes
dt	Deuteronomy	ho	Hosea	zc	Zechariah	lm	Lamentations
js	Joshua	jl	Joel	ma	Malachi	es	Esther
ju	Judges	am	Amos	1c	1 Chronicles	da	Daniel
1s	1 Samuel	ob	Obadiah	2c	2 Chronicles	er	Ezra
2s	2 Samuel	jn	Jonah	ps	Psalms	ne	Nehemiah
1k	1 Kings	mi	Micah	jb	Job		

The transliteration scheme used is known as the *Michigan-Claremont encoding scheme*, which originated in the work of a number of scholars at various universities in the mid-1980s. Transliteration was originally chosen for this text because at the time when this work was begun, it was not possible to display and manipulate both the Hebrew and Latin alphabet in a consistent way across computer systems. Even today, despite the widespread use of Unicode® font technology, there are still problems on many computer systems with displaying Hebrew accents, especially multiple accents on the same word.

The following table lists the transliteration system used in the WLC:

Consonants			Vowels		
Hebrew		Encoding	Hebrew		Encoding
Alef	א)	Patah	ַ	A
Bet	ב	B	Qamets	ָ	F
Gimel	ג	G	Segol	ֶ	E
Dalet	ד	D	Tsere	ֵ	"
Heh	ה	H	Hireq	ִ	I
Waw	ו	W	Holem	ׁ	O
Zayin	ז	Z	Qamets-Hatuf	ֱ	F
Het	ח	X	Qibbuts	׃	U
Tet	ט	+	Shureq	ׂ	W.
Yod	י	Y	Shewa	ְ	:
Kaf	כ ך	K	Hatef Patah	ֲ	:A
Lamed	ל	L	Hatef Segol	ֳ	:E
Mem	מ ם	M	Hatef Qamets	ִּ	:F
Nun	נ ן	N	Miscellaneous		
Samekh	ס	S	Dotless Sin/Shin		#
Ayin	ע	(Ketiv		*
Peh	פ ף	P	Qere		**
Tsade	צ ץ	C	Dagesh / Mappiq		.
Qof	ק	Q	Maqqef		-
Resh	ר	R	Raphe		,
Sin	שׁ שׂ	&	Accent (two-digit codes)		03 75 etc.
Shin	שׁ	\$			~
Taw	ת	T	Prefix/Suffix Separator		/

No distinction is made in our transliteration scheme between medial and final consonants, between *Qamets* and *Qamets-Hatuf*, between vocal and silent *Shewa*, or between *Dagesh forte* and *Dagesh lene* and *Mappiq*.

The forward slash character (“/”) was used by the original editors of this text to mark prefixes and suffixes on words. For example, the first word in the Hebrew Bible, <B. :/R")\$I75YT> is made up of the prefixed preposition <B. :> followed by the common noun <R")\$I75YT>. In Genesis 1:9, the word <L :/MIYN/080W> has both a prefix (the preposition <L :>) and a suffix (the pronominal suffix <080W>).

NOTE: In this document, transliterated Hebrew and Aramaic text will be enclosed between a less than symbol (“<”) and a greater than symbol (“>”). (The symbols “<” and “>” were chosen because they are not used for any characters in our transliteration system.)

Masoretic accents in the Hebrew Bible are represented by two-digit numbers. (For those who are interested in the origin of the accent codes and the transliteration system, see the included files **MICHIGAN.MAN** and **SUPPLMT.WTS**.)

Here are the two lines of Hebrew text from the WLC:

```
gn1:1 B. :/R")$I73YT B.FRF74) ):ELOHI92YM )"71T HA/$.FMA73YIM W:/)"71T HF/)F75REC00
gn1:2 W:/HF/)F81REC HFY:TF71H T033HW.03 WF/BO80HW. W:/X073$EK: (AL-P.:N"74Y T:H092WM
W:/R74W.XA ):ELOHI80YM M:RAXE73PET (AL-P.:N"71Y HA/M.F75YIM00
```

Hebrew/Aramaic accents occur after the consonant and vowel to which they belong, but before any *mater lectionis* consonant or the consonant that closes the syllable. (An exception to this rule is the case of the vowel *Shureq* (<W.>), which has accents after the *Waw* and *Dagesh*, as in the word <T033HW.03> in gn1:2, except when it is word-initial, when the accent will precede the *Waw*, as is seen in the word <14W./BFNFY/W> in gn7:7.)

Bracket Notes

One feature of the WLC is that we mark certain words with one or more special notes. For example, the first word of Genesis 2:10 is:

W:/NFHFR.03]p

This word is marked with the bracket note “]p” because we disagree with BHS on the punctuation of this word. We read an unexpected *Dagesh* in the *Resh* of this word in the Leningrad Codex, but BHS lacks the *Dagesh*. Thus we have bracket note to indicate that we disagree with BHS on the punctuation of this word.

When a bracket note applies to a word that ends with a *Maqqef*, the bracket note will appear after the word that follows that *Maqqef*. For example, in gn2:18 there is a bracket note “]1” on the two words <)E95(:E&EH.-L./071W]1>. This bracket note “]1” here happens to apply to the former word <)E95(:E&EH.-> and not to the latter word <L./071W>. (If there is a chain of three or more words separated by *Maqqefs*, all of the bracket notes applying to any of those words will follow the final word in the chain.)

We intend to publish on the Groves Center web site a document listing all of the bracket notes and the exact words to which they apply along with a specific explanation for each bracket note.

WLC 4.12 was the first release to include more than one bracket note on some words. In this release, we have 76 words (a total of 107 morphemes) with two bracket notes, 93 words (a total of 141 morphemes) with three bracket notes, four words (a total of five morphemes) with four bracket notes, and our first word (a word made up of two morphemes) with five bracket notes.

Sometimes we have more than one bracket note on the same word. For example, in Genesis 1:12 we find that the ninth and tenth words are:

(075&EH-P. :RI91Y]c]p

In this case, there is a “]c” bracket note on this pair of words because BHS has a *Mereka* (<71>) where we read a left *Meteg* (<75>) in the Leningrad Codex. We also see a *Maqqef* between these two words where BHS has a space. There are more than 116 occurrences of “]Q” followed by another bracket note in this release (4.20) where we agree with BHQ against BHS about the accents, consonants, vowels or punctuation in the text.

When two or more words are connected by *Maqqefs*, all bracket notes are placed after the last word. It is thus not clear to which word such bracket notes apply, but placing bracket notes immediately before *Maqqefs* was considered to more of a problem than a solution.

Another example of more than one bracket note on a word would be the seventh and eighth words of ca2:14:

)ET. -MAR:)A80Y/IK:]Q]p

For this occurrence, BHS lacks the *Dagesh* that is clearly present in the *Taw* in the Leningrad Codex, thus the “]p” to indicate that we disagree with BHS over punctuation and the “]Q” to indicate that we agree with BHQ. *When words are connected by one or more Maqqefs, the bracket notes are always placed after the last of those words.*

As of the WLC 4.14 release, we also have as many as four bracket notes on one word, such as the second word of ru3:14:

****MAR:G. :LOWTFY/OW03]Q]k]n]v**

In this case, we agree with BHQ about both the consonant *Waw* (the first occurrence, which BHS lacks) and the vowel *Holem* on the final syllable in this word (which BHS also lacks), thus the three bracket notes “]Q”, “]k” and “]v”, plus we have marked this unusual form with the new bracket note “]n” to flag it as anomalous (the *Holem* is unexpected with a *Qamets-Yod-Waw* pronomial suffix).

Another special case is the fourth word of er9:9:

W./B:/(AB:DU80T/"NW.]C]C]c

In this case, there are two “]c” and one “]c” bracket notes because we have two accent differences against BHQ and one accent difference against BHS 1997, respectively. The common difference that we have against both BHS and BHQ is that the *Zaqqef Parvum* is above the *Taw* in BHS 1997 and BHQ, instead of above the *Dalet* as in the Leningrad Codex and our text. The additional accent difference against BHQ is that BHQ has an *Azla* above the *Ayin* that is not found in the Leningrad Codex (or in BHS 1997, for that matter).

NOTES:

WLC 4.12 was the first release to include more than one bracket note on some words.

WLC 4.14 was the first release to include more than two bracket notes on some words. It was also the first release to include a repeated bracket note.

The 4.12 release included a new bracket note “]t” that marks words that differ from BHS in consonants, vowels or punctuation (spaces, *Sof Pasuqs*, *Maqqefs*, or *Dageshes / Mappiqs*). In WLC 4.14 and later, these “]t” bracket notes were *completely* replaced by the more specific bracket notes “]k”, “]v” or “]p” for consonants, vowels, or punctuation, respectively.

In WLC 4.8, we introduced a new bracket note “]c” to mark words in the morphology that differ from BHS in accentuation but not in the consonants, vowels, or punctuation. In the WLC 4.16 release and later, the bracket notes “]c”, “]k”, “]p”, and “]v” have completely replaced the old bracket note “]3”, which did not distinguish differences among accents, consonants, vowels, and punctuation.

The following chart lists the meanings of these bracket notes:

]1	BHS has been faithful to 𐤊 (the Leningrad Codex) where there might be a question of the validity of the form and we keep the same form as BHS. (This is similar to the note “]U ”, but the latter refers to cases where BHQ has been published and we keep the same form as both BHS and BHQ.)	gn32:24	L/OW00 (missing Silluq)
]2	WLC versions 4.4 and earlier inserted a <i>Sof Pasuq</i> at the end of the text of this verse because it would be expected, but those <i>Sof Pasuqs</i> were removed for version 4.6 through 4.14 because those <i>Sof Pasuqs</i> were not actually present in 𐤊. (The “]2 ” bracket notes were kept through WLC 4.14 for historical purposes, but have all been replaced in WLC 4.16 and later by “]1 ” (or, if appropriate, by “]U ”).)		(Only used in WLC 4.14 and earlier)
]3	We read or understand 𐤊 differently from BHS. Often this note indicates a typographical error in BHS. (“]3 ” was completely replaced by the more specific bracket notes “]c ”, “]k ”, “]p ”, and “]v ” in WLC 4.16 and later.)		(Only used in WLC 4.14 and earlier)
]4	<i>Puncta Extraordinaria</i> — a < 52 > is used to represent such marks in the text when they are above the line and < 53 > when below the line.	gn16:5	W./B"YNE75Y52/KF00
]5	Large letter(s)	lv11:42	G.FX061WN
]6	Small letter(s)	pr16:28	11W:/NIR:G.F81N
]7	Suspended letter(s)	ju18:30	M:NAS.E61H
]8	Inverted <i>Nun</i> (< N > in the text)	nu10:34	N
]9	BHS has abandoned 𐤊 and we (used to) concur. All of these occurrences are <i>Ketiv/Qere</i> problems. This bracket note became obsolete as of WLC 4.16. We changed the text at all of those places to follow 𐤊 and thus disagree with BHS.		(Only used in WLC 4.14 and earlier)
]C	We read an accent in 𐤊 differently from BHQ. (This is similar to the note “]c ”, but the latter refers to accent differences against BHS.)	dt28:25	YIT.EN/:KF63
]F	Marks a word with an anomalous consonant that is word-medial but has a word-final form.	is9:6	*L:/MAR:B."H
]K	We read a consonant in 𐤊 differently from BHQ. (This is similar to the note “]k ”, but the latter refers to consonant differences against BHS.)	er10:44	*NF&:)/IY This is the only occurrence of “]K ” in WLC 4.20 (as in WLC 4.16 and 4.18).

JM	Marks a word with an anomalous consonant that is word-final but has a word-medial form.	jb38:1	*MIN
JP	We read a punctuation character (<i>Maqqef</i> , <i>Mappiq</i> , <i>Dagesh</i> , space, or <i>Sof Pasuq</i>) in 𐤊 differently from BHQ. (This is similar to the note “ JP ”, but the latter refers to punctuation differences against BHS.)	lm3:14	& . :X0Q03
JQ	Marks a place where we agree with BHQ against BHS in reading 𐤊. This note will be preceded or followed by one or more bracket notes to specify the type(s) of disagreement(s) with BHS.	ru2:11) . FBI74Y/K :
JU	BHS and BHQ have both been faithful to 𐤊 (the Leningrad Codex) where there might be a question of the validity of the form and we keep the same form as both BHS and BHQ. This is similar to the note “ J1 ”, but the latter refers to cases where we keep the same form as just BHS. (“ JU ” only applies when the relevant fascicle of BHQ has been published and BHQ has the same reading as BHS.)	dt23:18	YI& : RF)" L0 (missing Silluq)
JV	We read a vowel in 𐤊 differently from BHQ. (This is similar to the note “ Jv ”, but the latter refers to vowel differences against BHS 1997.)	es1:14) "LF81Y/W
Ja	Adaptations to a <i>Qere</i> that 𐤊 and BHS, by their design, do not indicate. Usually this indicates the addition of a <i>Maqqef</i> to our <i>Qere</i> text that is not present in the margin of 𐤊, or to the addition of a <i>Dagesh</i> or <i>Mappiq</i> to our <i>Qere</i> text that is not present in the <i>Ketiv</i> consonants in the main text of 𐤊.	ex4:2 je50:29 lm4:16	**MAH- (The <i>Maqqef</i> in the <i>Qere</i> is not present in the margin of 𐤊.) **L/FH.03 (The <i>Mappiq</i> in the <i>Qere</i> is not present in the <i>Ketiv</i> text of 𐤊.) **W./Z:Q"NI73YM (The <i>Dagesh</i> in the <i>Waw</i> is not present in the <i>Ketiv</i> text of 𐤊 because there is no <i>Waw</i> in the <i>Ketiv</i> text.)
Jc	We read an accent in 𐤊 differently from BHS. (This is similar to the note “ Jc ”, but the latter refers to accent differences against BHQ.)	gn22:2	B. IN/ :KF63

]k	We read a consonant in 𐤊 differently from BHS. (This is similar to the note “]K ”, but the latter refers to consonant differences against BHQ.)	gn14:2	**C : BOWYI80YM
]m	Miscellaneous notes to the text and occasions where more than one bracket category applies. All of these notes were re-examined for WLC 4.16 and either removed or replaced with a more appropriate bracket note.		(Only used in WLC 4.14 and earlier)
]n	An anomalous form in the text of 𐤊. This bracket note was new to WLC 4.14 (only two occurrences in that version), and was not yet widely used at first, occurring on only 23 words in WLC 4.16. However, this bracket note occurs on 111 words (167 morphemes) in WLC 4.20.	lv25:20	N . O) KA7073L (In this case, the double accent is anomalous.)
]p	We read a punctuation character (<i>Maqqef</i> , <i>Mappiq</i> , <i>Dagesh</i> , space, or <i>Sof Pasuq</i>) in 𐤊 differently from BHS. (This is similar to the note “]P ”, but the latter refers to a punctuation difference against BHQ.)	gn2:10	W : /NFHFR . 03]p
]q	We have abandoned or added a <i>Ketiv/Qere</i> relative to BHS. In doing this we agree with 𐤊 against BHS.	gn9:21 ps21:2)FH : FL / 075H00 (Here we do <i>not</i> have a <i>Qere</i> reading where BHS <i>does</i> have one.) **Y . F71GEL (Here we <i>do</i> have a <i>Qere</i> reading where BHS does <i>not</i> have one.)
]t	We read one or more consonants, vowels or punctuation (<i>Maqqef</i> , <i>Dagesh</i> , <i>Mappiq</i> , or <i>Sof Pasuq</i>) in 𐤊 differently from BHS. (This bracket note was only used in WLC 4.12.)		(This note was only used in WLC 4.12 and was replaced in WLC 4.14 and later by the more specific bracket notes “]c ”, “]k ”, “]p ”, and “]v ”.)
]v	We read a vowel in 𐤊 differently from BHS. (This is similar to the note “]V ”, but the latter refers to vowel differences against BHQ.)	gn31:42	Y : GI94Y (A
]y	<i>Yathir</i> readings in 𐤊 which we have designated as <i>Qeres</i> . In WLC 4.14 and earlier, a few <i>Yathir</i> readings in 𐤊 were not marked as such because they were not treated as <i>Qeres</i> in either BHS or in the Dotan ADI edition of 1973, but beginning with 4.16 we rely only on whether or not something is a <i>Yathir</i> reading in 𐤊.	1s28:8	**QF95S : FMIY - NF71)

The Ketiv-Qere

This Masoretic feature presents more challenges to the data structure than any other single factor. First of all, we must list the *Ketiv* and *Qere* as separate words (unlike BHS which uses one graphical word plus consonants in the margin to indicate both). Thus word counts in verses with *Ketiv* / *Qere* will be skewed by the number of *Ketiv* / *Qere* in a verse.

Secondly, there are simply a number of anomalous situations created by trying to account for this feature within our normal encoding. A list of the kinds of situations follows.

1. Simple *Ketiv-Qere*

gn8:17 *HWC) **HAY:C"74)

2. *Ketiv* following *Maqqef*

2s3:25 W:/)ET-*MBW)/K **M074WBF)/E80KF W:/LF/DA85(AT

3. *Ketiv* and *Qere* with *Maqqef*

2s14:7 L:/BIL:T.I94Y *#WM- **&IYM-L:/)IY\$/I91Y

4. *Ketiv* and *Qere* internal to compound name

2k23:10 B.:/G"74Y *BNY- **BEN-HIN.092M

(js18:24 and 2k14:7 have the only such other occurrences.)

5. *Qere* without *Ketiv* (*Qere wela Ketiv* - “qwlk”)

2s8:3 B.I75/N:HAR- *kk **P.:RF75T00]a

8. *Ketiv* without *Qere* (*Ketiv wela Qere* - “kwlq”)

ek48:16 *XM# **qq

Coding Practices

1. Slash (/) to indicate the end of a prefix or the beginning of suffix in textual entry

A slash in the text indicates the end of a prefix or the beginning of the suffix.

2s2:5 WA/Y.I\$:LA70X . . .):AL"Y/HE81M

The slash after the <WA> indicates that the <WA> is a prefix. The slash before the <HE81M> indicates that the <HE81M> is a suffix.

2. Directional suffix on first word of a compound name

(Potential confusion because directional-*He* (or *He*-locale) is suffixed to the first word in the compound and not the second – “/” indicates suffix position.)

gn28:2 P.AD.E74N/F75H~):ARF80M

3. Two suffixes on a single word (verbs only)

Several times in Isaiah, Jeremiah, Hosea, Psalms, and Proverbs, there is both a paragogic *Nun* and an object suffix which means an analysis will indicate two suffixes. For example:

ps63:4 Y:\$AB.:X75W./N/:KF0

4. Article as part of proper names (“//”)

In early versions of the WLC, a double slash was used to indicate a definite article as part of a proper name, but this practice was discontinued in WLC 4.4. For example, in WLC 4.2 we had:

gn35:27 . . . QIR:YA74T HF75//)AR:B.A92(. . .

but in WLC 4.4 through 4.20 this same proper name appears as:

gn35:27 . . . QIR:YA74T HF75)AR:B.A92(. . .

5. Run-on words (textual corruption in manuscript)

Run-on words, a textual critical issue, must be analyzed separately. In such cases there is no space or *Maqqef* in the text as would be expected, so they must be treated as a single orthographic entity, though we have added one (or more) morphological slash(es) to separate the distinct morphemes. For example:

ps106:1 HA75L:LW./YF63H.05

There are other examples of run-on words in is9:5 (word 18) and ec4:10 (word 8) and also in *Ketiv* text in 2k6:25 (word 16), ek8:6 (word 7), er4:9 (word 17), and 1c9:4 (word 9). (The run-on *Ketiv* word in ec8:6 was introduced in WLC 4.18.)

6. *Raphe*

Biblia Hebraica Stuttgartensia's editorial policy was to omit *Raphe* consistently, "...otherwise there would have been almost insuperable technical difficulties."¹ Nevertheless, *Raphe* does occur in BHS twelve times. We have replicated these cases in WLC, where *Raphe* is represented in the encoded Hebrew text as a comma (","):

```
ex20:13 T., IR:CF7375X00
ex20:14 T., IN:)F9275P00
ex20:15 T., IG:N08075B00
dt5:13 K., FL-M:LA)K:T./E75KF0100
dt5:17 T., IR:CF75X00
dt5:18 T., IN:)F7592P00
dt5:19 T., IG:N07580B00
2s11:1 M.AL:),KI81YM]1
is22:10 T.IT,:CW.03
je20:17 RAX:MF73H,
zc5:11 L/F71H,
ps119:99 L,/I75Y00
```

These occurrences are almost all textual issues, most of them occurring in the Decalogues of Exodus and Deuteronomy where there is the well-known double punctuation. As a courtesy to users, WLC has the same *Raphes* as BHS. The Leningrad Codex uses *Raphe* extensively, and adding those to the Hebrew text of WLC is planned for a future release.

Additional documentation is available in the enclosed files "**MICHIGAN.MAN**" and "**SUPPLMT.WTS**".

¹ Karl Elliger and Wilhelm Rudolph, "Foreword" in *Biblia Hebraica Stuttgartensia*, p. xii.