The Unicode/XML Leningrad Codex

The Westminster Leningrad Codex (*WLC*) began as an electronic transcription by Richard Whitaker (Princeton Seminary, New Jersey) and H. van Parunak (then at the University of Michigan, Ann Arbor) of the 1983 printed edition of *Biblia Hebraica Stuttgartensia* (*BHS*). It was continued with the cooperation of Robert Kraft (University of Pennsylvania) and Emmanuel Tov (Hebrew University, Jerusalem), and completed by Prof. Alan Groves. The transcription was called the *Michigan-Claremont-Westminster Electronic Hebrew Bible* and was archived at the Oxford Text Archive (OTA) in 1987. It has been variously known as the "CCAT" or "e-BHS" text. Since that time, the text has been modified in many hundreds of places to conform to the photo-facsimile of the Leningrad Codex, Firkovich B19A, residing at the Russian National Library, St. Petersburg; hence, the change of name to the Westminster Leningrad Codex.

In 2016, Christopher V. Kimball first published a fork of WLC Version 4.20, transferred to a more modern (Unicode/XML) format; that text is updated semi-annually to better match the *Leningrad Codex* from reader suggestions, through a formal and automated process – more than a thousand changes have been made to date. The Hebrew text is version controlled to provide a fixed reference for derived work. The text in this collection is that of "Build 27.3, 18-Oct-2024" – updates will be made periodically.

The Unicode/XML Leningrad Codex is in the public domain and is included here with permission.

The 'TheoDoor' Version

The text here presented is reproduced from the text archive of the **Unicode/XML Leningrad Codex** described above; the provenance of this text, summarized from release notes in various headers, is shown below:

####	Unicode/XML Leningrad Codex [UXLC 1.2]
####	Build: 26.2 – 11 Mar 2021 06:46
####	Publisher: Christopher V. Kimball, West Redding, CT USA
####	From WLC 4.20 provided by the Groves Center.
####	WLC2XML [18 Dec 2020] – Transcription from Michigan-Claremont encoding to Unicode with XML markings.
####	Unicode/XML Leningrad Codex [UXLC 2.2]
####	Build: 27.3 – 18 Oct 2024 05:59
####	Layout: Full; Content: Accents.

Each chapter is placed in a separate section; the font used is Ezra SIL (a font designed specifically to match the printed *BHS*; 14-point for the body text and 16-point for the chapter titles); verse numbers have been marked as orange superscript and a (single) non-breaking space character kept after each. We have also marked 'special characters' (enlarged, small, or raised) according to the *BHS* transcription, included 'true' paragraph breaks after *Petuhah* (open space) markers, and highlighted the *Setumah* and *Petuhah* markers with a different (orange) font color.

Unlike most texts in this collection, which are saved as Word documents ('docx' files), the *UXLC* texts are saved as OpenDocument Text ('odt' files). This is because – for reasons yet to be resolved – these texts do not lend themselves well to being portable when saved as Word files (for example, LibreOffice will hang when trying to open the larger files).

Genesis	1 Samuel	Esther	Lamentations	Micah
Exodus	2 Samuel	Job	Ezekiel	Nahum
Leviticus	1 Kings	Psalms	Daniel	Habakkuk
Numbers	2 Kings	Proverbs	Hosea	Zephaniah
Deuteronomy	1 Chronicles	Ecclesiastes	Joel	Haggai
Joshua	2 Chronicles	Song of Solomon	Amos	Zechariah
Judges	Ezra	Isaiah	Obadiah	Malachi
Ruth	Nehemiah	Jeremiah	Jonah	