To: Unicode Script Encoding Working Group
Author: Ben Denckla
Subject: Adding *dagesh ḥazaq* to Hebrew
Date: 28 May 2025

I propose that a *dagesh ḥazaq* code point be added to the Hebrew block. Unicode could then make graphical distinctions that are needed in some types of Hebrew publishing. Also, Unicode could then make semantic distinctions that are needed in some applications like phonetic transcription and grammatical analysis.

1. Background

The existing code point U+05BC (Hebrew Point *Dagesh* or *Mapiq*) encodes four dots that are semantically distinct, though traditionally they have been graphically indistinct:

1. *Dagesh ḥazaq* (דגש חזק). This dot appears in most but not all letters.

2. *Dagesh qal* (דגש קל). This dot appears only in the six letters בגדכפת and their final forms, ך and ף. *Dagesh ḥazaq* can appear in most of these letters, so it must be distinguished from *dagesh qal* by context or graphically.

3. *Shuruq* dot (נקודת שורוק). This dot appears only in the letter *vav* (ו), making that *vav* a *shuruq*.[1] *Dagesh ḥazaq* can appear in *vav* so it must be distinguished from *shuruq* dot by context or graphically.

4. *Mapiq* (מפיק). This dot appears at least in the letters *he* (ה) and *alef* (א).[2] None of the other three dots appear in *alef* or *he*. So, none of them need to be distinguished from *mapiq* in *alef* or *he*. The topic of *mapiq* in *vav* (ו) and *yod* (י) is beyond our scope here.[3]

In recent decades, many Hebrew publications have made the following distinctions:

1. One shape for *dagesh qal*. This shape is also used for *shuruq* dot and *mapiq*.

2. Another shape dedicated to *dagesh ḥazaq*. This shape is not used for any other purpose.

---

[1] Unicode's annotation "= *shuruq*" for U+05BC should really say "= *shuruq* **dot**" because a *shuruq* is a *vav* qualified by this dot; the dot alone is not called *shuruq*.

[2] Khan argues that the standard four dots in א in *Tanakh* are *dagesh ḥazaq* dots not *mapiq* dots. Our main concern is what distinctions need to be made in each letter, and either way, no distinctions need to be made in *alef*.

[3] Some nonstandard manuscripts have clear uses of *mapiq* on both *vav* and *yod*. See Gesenius 14d, 8m and Khan I.1.10. But the standard manuscripts and the printed tradition have, at most, some unclear uses of *mapiq* on *vav*. Those cases have a *vav* with a dot that can be interpreted as a *mapiq*, as in Exodus 35:26 טָו֖וּ. But the interpretation of this dot is unclear. See Yeivin #396 and Dotan xiv.

Though only recent publications make these distinctions graphically, these distinctions have been well-established, semantically, for hundreds of years in the theory of the grammar and phonetics of the Hebrew language.

2. Proposed characters

I propose a new code point called HEBREW POINT DAGESH HAZAQ. I propose the following annotations for the existing and new code points (trying to mimic the wording of the annotation currently used for U+05B8 (*Qamats*)):

U+05BC Hebrew Point *Dagesh* or *Mapiq*

> = *shuruq* dot

> • falls within the base letter; as *dagesh*, used either generically or as *dagesh qal* in texts that distinguish it from *dagesh hazaq*

> → U+05C9 Hebrew Point *Dagesh Hazaq*

U+05C9 Hebrew Point *Dagesh Hazaq*

> = *dagesh forte*

> • used in texts that distinguish *dagesh qal* from *dagesh hazaq*

> → U+05BC Hebrew Point *Dagesh* or *Mapiq*

The suggestion above assumes that the new code point is assigned location U+05C9, but the exact location is of course not important.

The suggestion above assumes that the annotation for U+05BC (*Dagesh* or *Mapiq*) can be updated, as seems to have been the case for U+05B8 (*Qamats*) when U+05C7 (*Qamats Qatan*) was introduced.

Proposing only one new code point, for *dagesh ḥazaq*, implicitly proposes that the existing code point, U+05BC, would then do "double duty":

- In some texts, U+05BC would continue to represent all four dots, just as it currently does.

- In other texts, U+05BC would start to represent only the three dots *dagesh qal*, *mapiq*, and *shuruq* dot, while the new code point would represent only *dagesh ḥazaq*.

Despite its ambiguity, this "double duty" seems to be the preferred way to disunify a diacritic mark in the Hebrew block, judging from the way the following code points were added:

- U+05C7 (*Qamats Qatan*), added in Unicode 4.1: no *qamats gadol* code point was added. Instead, the existing U+05B8 (*Qamats*) took on "double duty" once U+05C7 was added.

- U+05BA (*Holam Haser* for *Vav*), added in Unicode 5.0: no "*ḥolam malei* for *vav*" code point was added. Instead, the existing U+05B9 (*Holam*) took on "double duty" once U+05BA was added.

This "double duty" is also consistent with the pending proposal for a *sheva na* (mobile *sheva*) code point, [L2/24-274](): no *sheva naḥ* (resting *sheva*) code point is proposed. Instead, the existing U+05B0 (*Sheva*) will take on "double duty" if and when U+05C8 is added.

3. Properties

The properties of the new code point would be the same as the existing code point U+05BC (*Dagesh* or *Mapiq*).

4. Collation

The collation of the new code point would be the same as the existing code point U+05BC (*Dagesh* or *Mapiq*).

5. References

Dotan, Aaron. *Biblia Hebraica Leningradensia*. United States: Hendrickson Publishers, 2001.

Gesenius, Wilhelm and Cowley, Arthur Ernest. *Gesenius' Hebrew Grammar*. United Kingdom: Clarendon Press, 1910.

Khan, Geoffrey. *The Tiberian Pronunciation Tradition of Biblical Hebrew*. United Kingdom: Open Book Publishers, 2020.

Yeivin, Israel and Revell, E. J. *Introduction to the Tiberian Masorah*. United States: Scholars Press, 1980.

6. Acknowledgments

7. Examples

Below I provide examples from the following 5 publishers using *dagesh ḥazaq*:

1. פלדהיים (Feldheim) (paper)
2. מנהגי אבות (*Minhagei Avot*) (paper)
3. מכון הרב מצליח (The Rav Matzliach Institute) (paper)
4. למען שמו באהבה (*Lema·an Shemo Be·ahavah*) (paper)
5. על־התורה (Al-Hatorah) (web)

Feldheim is one major publisher making these graphical distinctions. For example, here is a picture showing both types of *dagesh* in the word תֹּאכְלֶ֫נָּה ("you shall eat of it") in Genesis 3:17 in the *Tanakh* (Hebrew Bible) made under Feldheim's Simanim imprint:

dagesh ḥazaq   dagesh qal

Here is a picture showing both *dagesh ḥazaq* and *shuruq* dot in the word וַיְצַוּוּ ("and charged") in Joshua 3:3 in the same Simanim publication:
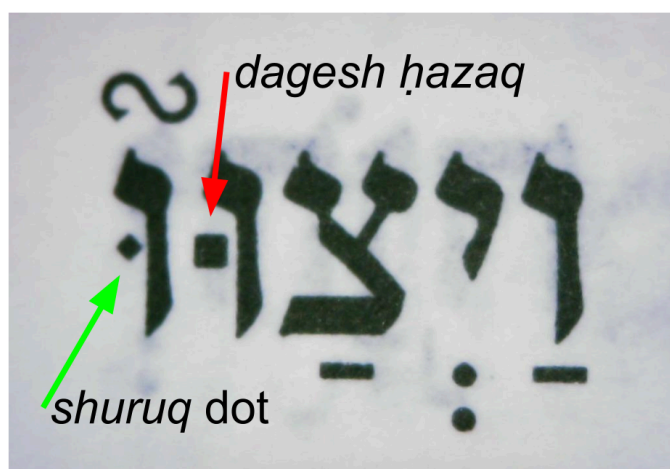
The publisher *Minhagei Avot*, particularly in its *Darkei Avot* (דרכי אבות) series, also makes these graphical distinctions. Here is a picture showing both types of *dagesh* in the word הַמִּקְדָּשׁ ("the sanctuary") in the *Darkei Avot* prayerbook (*siddur*):

The Rav Matzliach Institute (מכון הרב מצליח) is a publisher using a distinctive *dagesh ḥazaq* in a wide variety of publications including prayerbooks (*siddurim*), editions of the Psalms (*Tehillim*), and editions of the Pentateuch (Five Books of Moses) (*Torah*). Here are a few words, מַעַלְלֵי־יָהּ כִּי־אֶזְכְּרָה מִקֶּדֶם ("deeds of Yah that I recall; from days of old"), from Psalm 77:12 in one of its editions of Psalms:

The Rav Matzliach Institute considers their *dagesh ḥazaq* to be an important enough feature to be listed on most or all of the title pages of their relevant books. For example here is an image of the title page of their *ḥumash* (a type of edition of the Pentateuch), prominently listing their *dagesh ḥazaq* feature (along with their *sheva na* and *qamats ḥatuf* (*qamats qatan*) features):

Another important publisher featuring *dagesh ḥazaq* (along with *sheva na* and *qamats qatan*) is *Lema·an Shemo Be·ahavah* (למען שמו באהבה). Their prayerbooks (*siddurim*) and editions of the Pentateuch (*ḥumashim*) sell in high volume in Israel, and most or all include these special marks. Like the Rav Matzliach Institute, they advertise their use of these special marks prominently on most or all of the title pages of their relevant books. Here is the word בַּבְּהֵמָה ("with [every] beast") from Genesis 9:10 in a *ḥumash* of theirs:

These distinctions are not just present in paper publications. Despite great technical challenges due to lack of Unicode support, these distinctions are present in the *Tanakh* of [Al-Hatorah's Mikraot Gedolot](#) website. For example, here is an image showing both types of *dagesh* in the chanted word בְּתוֹךְ־הַגָּן ("in the middle of the garden") in Genesis 3:3 in the Al-Hatorah *Tanakh*:

8. ISO Proposal Summary

(The ISO proposal summary forms will appear here but the Template for Character Additions advises that these need not appear here in this preliminary version of the proposal.)

9. Supporters

(List of supporters (people who support this idea, not financial supporters) will be added here.)

(End of document.)

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Addition of Hebrew Point Dagesh Hazaq*
2. Requester's name: *Ben Denckla*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2025-07-15*
5. Requester's reference (if applicable):
6. Choose one of the following:
   This is a complete proposal: *Yes*
   (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): *No*
      Proposed name of script:
   b. The proposal is for addition of character(s) to an existing block: *Yes*
      Name of the existing block: *Hebrew*
2. Number of characters in proposal: *1*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   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 ☐
4. Is a repertoire including character names provided? *Yes*
   a. If YES, are the names in accordance with the "character naming guidelines"
      in Annex L of P&P document? *Yes*
   b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
   a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
      Ben Denckla
   b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
      Ben Denckla, Ben.Denckla@GMail.com
6. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources)
      of proposed characters attached? *Yes*
7. Special encoding issues:
   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)? *No*

8. Additional Information:

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/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

---

[1] Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?   *No*
     If YES explain

2. 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*
         If YES, with whom?   *See list of supporters in proposal*
         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?   *Yes*
     Reference: *See proposal*

4. The context of use for the proposed characters (type of use; common or rare)
     Reference: *See proposal*

5. Are the proposed characters in current use by the user community?
     If YES, where? Reference: *Yes, see proposal*

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
     in the BMP?   *Yes*
         If YES, is a rationale provided?   *No*
         If YES, reference:

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

8. Can any of the proposed characters be considered a presentation form of an existing
     character or character sequence?   *No*
         If YES, is a rationale for its inclusion provided?
         If YES, reference:

9. Can any of the proposed characters be encoded using a composed character sequence of either
     existing characters or other proposed characters?   *No*
         If YES, is a rationale for its inclusion provided?
         If YES, reference:

10. Can any of the proposed character(s) be considered to be similar (in appearance or function)
     to, or could be confused with, an existing character?   *Yes*
         If YES, is a rationale for its inclusion provided?   *Yes*
         If YES, reference: *See proposal*

11. Does the proposal include use of combining characters and/or use of composite sequences?   *No*
     If YES, is a rationale for such use provided?
         If YES, reference:
     Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
         If YES, reference:

12. Does the proposal contain characters with any special properties such as
     control function or similar semantics?   *No*
         If YES, describe in detail (include attachment if necessary)

13. Does the proposal contain any Ideographic compatibility characters?   *No*
     If YES, are the equivalent corresponding unified ideographic characters identified?
         If YES, reference: