Doc Type: Working Group Document
Title: Proposal to add ATNAH HAFUKH to the BMP of the UCS
Source: Mark Shoulson, Peter Kirk, and Michael Everson
Status: Individual Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Date: 2003-12-11

This document requests an additional character to be added to the UCS and contains the proposal summary form.

A. Administrative

1. Title
Proposal to add ATNAH HAFUKH to the BMP of the UCS.
2. Requester's name
Mark Shoulson, Peter Kirk, and Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2003-12-11
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.
Proposed name of script
1b. The proposal is for addition of character(s) to an existing block
Yes.
1b. Name of the existing block
Hebrew
2. Number of characters in proposal
1
3. Proposed category (see section II, Character Categories)
Category B.1
4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)
Level 3.
4b. Is a rationale provided for the choice?
Yes.
4c. If YES, reference
Combining character.
5a. Is a repertoire including character names provided?
Yes.
5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?
Yes.

5c. Are the character shapes attached in a legible form suitable for review?
Yes.

6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
Michael Everson. TrueType.

6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
Michael Everson. Fontographer.

7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
Yes, see bibliography below.

7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?
Yes.

8. 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, see below.

9. 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.
Yes, see Unicode properties below.

C. Technical – Justification

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

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?
Professor Aron Dotan, script expert

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?
Yes.

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

4b. Reference
See examples below.

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

5b. If YES, where?
See examples below.

6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?
Yes.

6b. If YES, is a rationale provided?
Yes.

6c. If YES, reference
All Hebrew cantillation marks are in the BMP.

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

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?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference
10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
No.
10b. If YES, is a rationale for its inclusion provided?
10c. If YES, reference
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)?
Yes.
11b. If YES, is a rationale for such use provided?
No.
11c. If YES, reference
12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
No.
12b. If YES, reference
13a. Does the proposal contain characters with any special properties such as control function or similar semantics?
No.
13b. If YES, describe in detail (include attachment if necessary)
14a. Does the proposal contain any Ideographic compatibility character(s)?
No.
14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?
14c. If YES, reference

D. Proposal

The Hebrew Bible is traditionally annotated with cantillations or accents, which indicate a melody for chanting the text and also serve to mark syntactic units. Two distinct systems are used in the text (though they clearly influence one another): one (sometimes called the "Poetic System") in the "3 Books" (Job, Proverbs, and Psalms), and one (the "Prose System") in the remaining 21 Books of the Hebrew Bible, as traditionally divided. One of the cantillation accents in use in the Masoretic text of the Hebrew Bible, U+05AA HEBREW ACCENT YERAH BEN YOMO, occurs rarely in the Prose system, serving as a connecting accent before U+059F HEBREW ACCENT QARNEY PARA (which is also called *pazer gadol* or 'great *pazer*'). It occurs rather more frequently in the Poetic System, where it appears before U+05A1 HEBREW ACCENT PAZER and also before the accent *ole v'yored* (a composite accent, consisting of U+05AB HEBREW ACCENT OLE and U+05A5 HEBREW ACCENT MERKHA). Both uses of this accent (before *qarney para*/*pazer* and before *ole v'yored*) have been printed with the same glyph for the last few centuries, which is why they had been unified. But research of the past few decades has revealed that in fact they are not the same accent, and some scholars and printers have recognized that they must make the distinction explicit.

Examination of ancient Bible manuscripts and early printed editions shows that the accent used before both *pazer*s (i.e. U+059F HEBREW ACCENT QARNEY PARA, the great *pazer* in the Prose System, and U+05A1 HEBREW ACCENT PAZER, the lesser *pazer*, which occurs in both systems but is accompanied by the accent under discussion only in the Poetic System) is distinct from the accent used before *ole-v'yored*. For example, in the Leningrad Codex, scans of which are attached, the accent used before both *pazer*s was written with a rounded glyph, like the bottom three-fourths of a circle (which matches the names used: *galgal* 'wheel' and *yerah ben yomo* 'day-old moon'). However, the accent used in the Poetic System in conjunction with the *ole-v'yored* was written with a pointed glyph, like a small letter "v", or an inverted form of U+0591 HEBREW ACCENT ETNAHTA as written in those manuscripts.

We therefore suggest that the sample glyph for U+05AA HEBREW ACCENT YERAH BEN YOMO be changed to a stemless simple curved form, and the proposed HEBREW ACCENT ATNAH HAFUKH be given either the glyph currently used for U+05AA or a "v"-shaped glyph. This distinction was noticed by Yisrael Yeivin, who named the accent before the *ole v'yored* "*atnah hafukh*" or "inverted *etnahta*", after its form. It is

now in use in several printed works (see examples below), including the Bible used by the Israeli Defense Forces, edited by Aron Dotan, and the "Simanim" publication of the book of Psalms.

| 05AA | ◌ | HEBREW ACCENT YERAH BEN YOMO (glyph change) |
| 05A2 | ◌ | HEBREW ACCENT ATNAH HAFUKH |

Unicode Character Properties

```
05A2;HEBREW ACCENT ATNAH HAFUKH;Mn;220;NSM;;;;;N;;*;;;
```

Bibliography

Breuer, M., טעמי המקרא בכ״א ספרים ובספרי אמ״ת, Horev Press, Jerusalem, 1989.
Dotan, Aron, ed., תורה נביאים כתובים לחיילי צבא־הגנה לישראל, 1982.
Riahi, S. M., ed., "סימנים" ספר תהלים, Feldheim Press, Jerusalem, 2001.

Figures

We will begin with some examples of modern texts which distinguish the two cantillations.
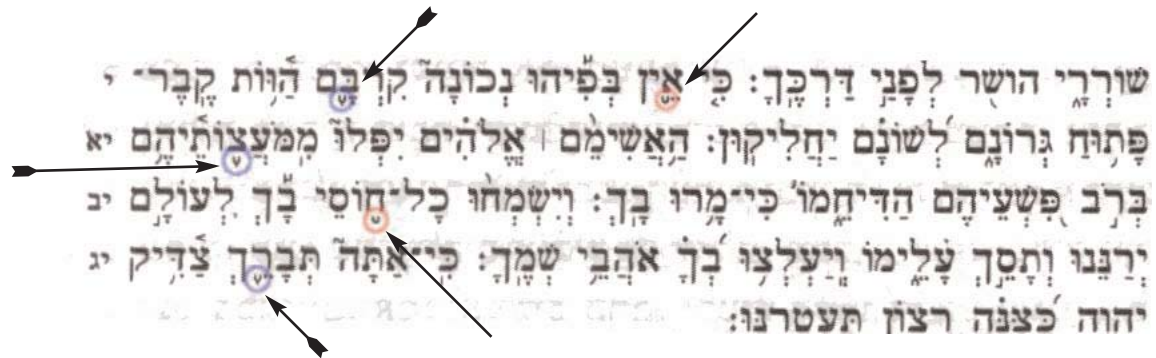
Figure 1. From the IDF Bible (Psalms 5:10-13), edited by Aron Dotan. In this picture and those that follow, instances of U+05AA HEBREW ACCENT YERAH BEN YOMO (i.e. the accent before *pazer*) will be circled in red and pointed to with a plain arrow, and instances of the proposed HEBREW ACCENT ATNAH HAFUKH will be circled in blue and pointed to with a feathered arrow.
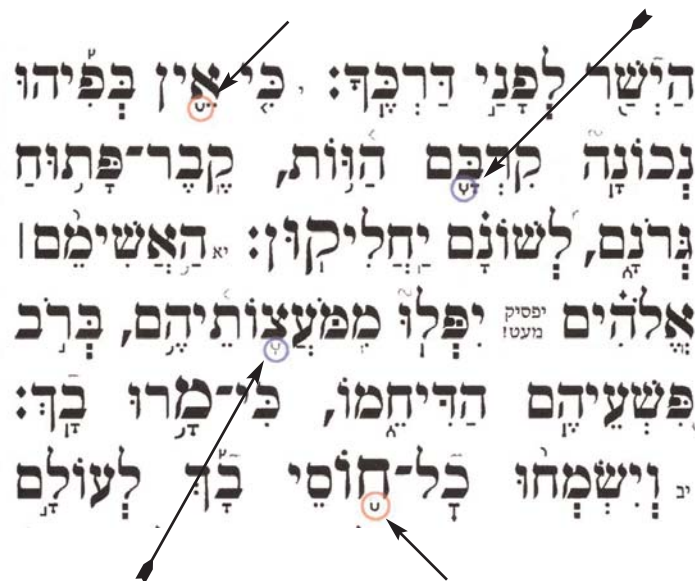

Figure 2. From the "Simanim" Psalms.

1. מֵרְכָא –
2. מֻנַּח –
3. עִלּוּי –
4. מַרְחָא –
5. גַּלְגַּל *[גַּלְגַּל]–
6. מְהֻפָּךְ –
7. אַזְלָא –

וּבְתוֹרָתוֹ

וּבְדֶרֶךְ

יִתֵּן

וְכָל

אֹמְרִים

אֲשֶׁר

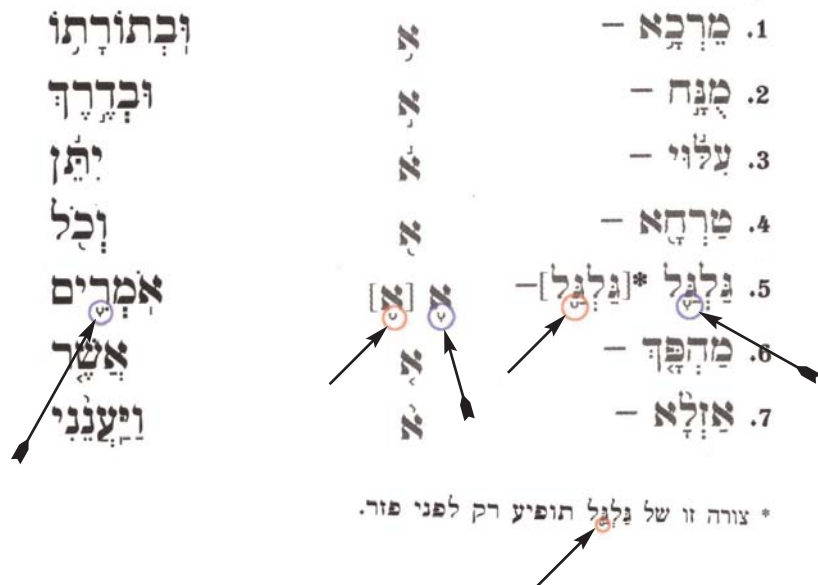וַיַּעֲנֵנִי

* צורה זו של גַּלְגַּל תופיע רק לפני פזר.

Figure 3. From the same book, part of a list of the cantillations and their forms and names. Note the two different forms given for the *galgal*. The sentence marked with the asterisk reads: 'This form of *galgal* will appear only before a *pazer*.'
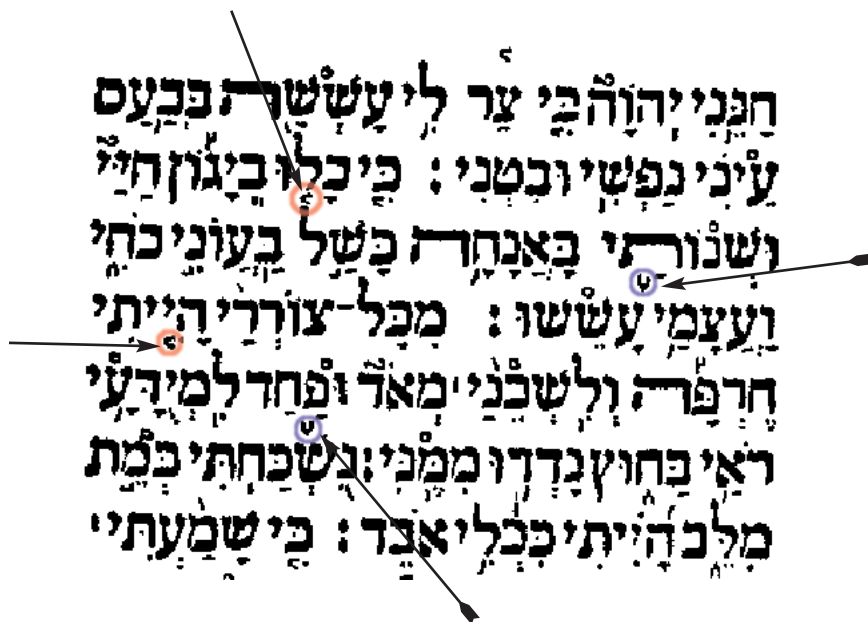

Figure 4. From the "Rabbinic Bible", printed in 1524-1525 CE. Here the sources cited are a little less modern. Note that the YERAH BEN YOMO has a symbol rather different from any we have seen so far (looks a little like a C), while the ATNAH HAFUKH is definitely distinct from it, with an upside-down ETNAHTA form.
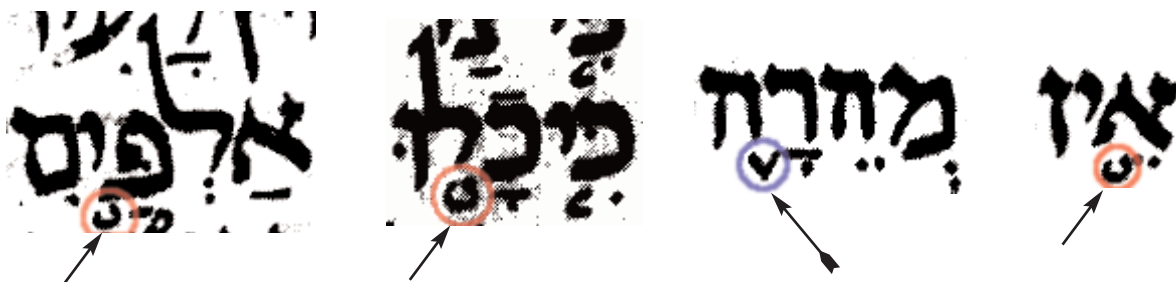

Figure 5. Some examples from the Leningrad Codex. Note the distinction made in the glyphs. The leftmost example is from Numbers, showing that the rounded form is the one used in the "21 Books".
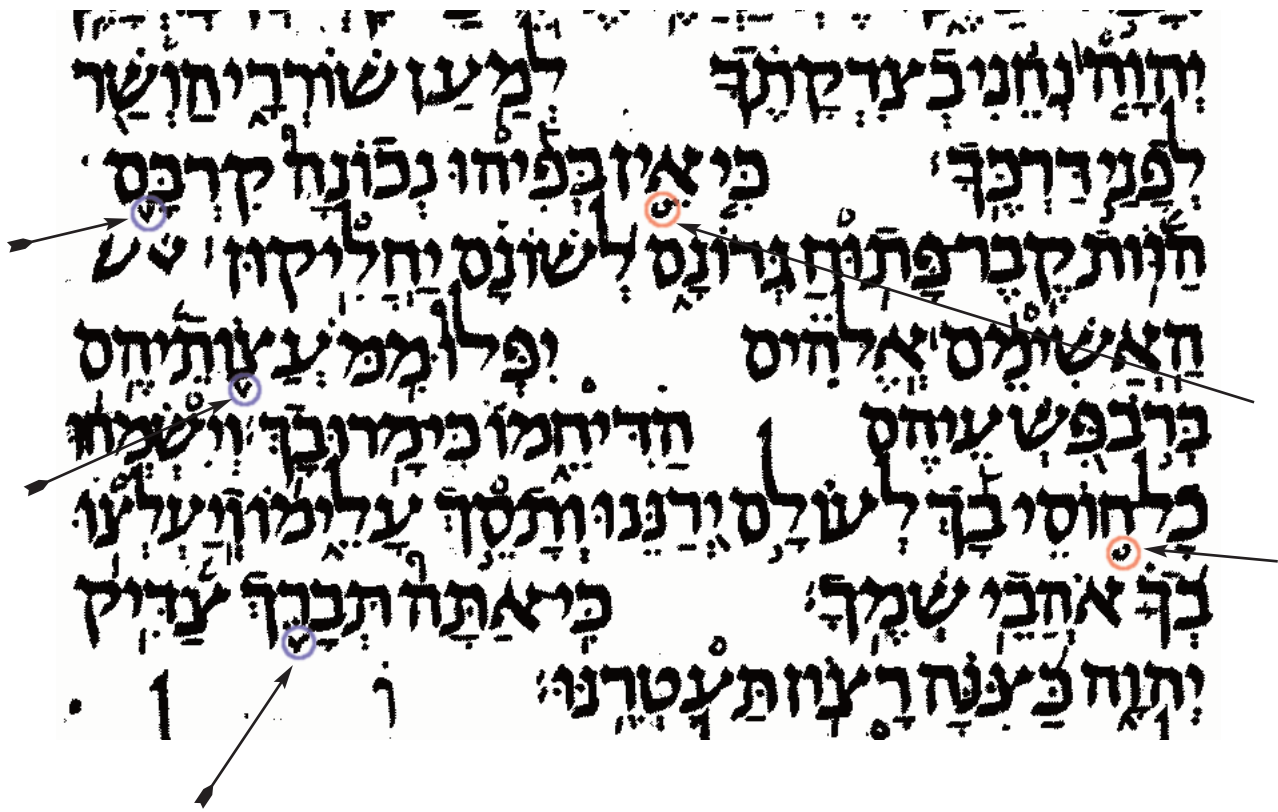
Figure 6. A larger example from the Leningrad Codex, showing several examples of both marks, all clearly distinct.
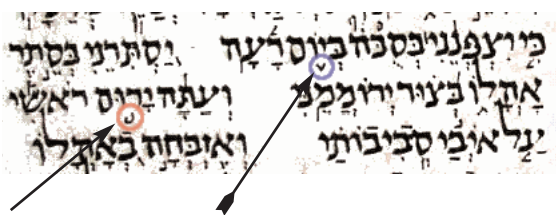

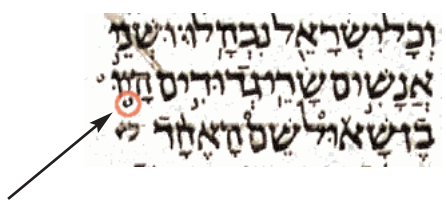**Figure 7.** An example from the Aleppo Codex, showing both forms.

Figure 8. Another example from the Aleppo Codex, this one from 2 Samuel, demonstrating that the YERAH BEN YOMO in the "21 Books" is in fact the rounded form.