To: Unicode Script Encoding Working Group (SEWG)
Author: Ben Denckla
Subject: Adding Hebrew stress helper accents
Date/Time: 16 September 2025 (23 Elul 5785)

I propose adding code points for six Hebrew stress helper accents.

| | |
|---|---|
| In about 7,800 words in the Hebrew Bible, the primary stress is not marked by the primary accent. It may be marked by a helper accent, or it may not be marked at all. For example, to the right we show three versions of Gen. 3:17 אשתך. Each version shows, in orange or green, a syllable that might have the stress. It happens to be the green one, but we have no way of knowing that with just a *segol* accent at the end of the word (shown with a blue arrow in the first version). |  |
| In some texts (manuscripts and printed editions), this word's stress is marked by a *segol* helper above the *tav* (ת), as shown to the right. |  |
| Only some texts use helpers for *segol*. But almost all texts consistently use helpers for *pashta*, as shown to the right in Gen. 1:2 תהו. Of the 7,800 or so words that need a helper, about 3,800 need a *pashta* helper and almost always get it. How many of the remaining 4,000 or so words get the helper they need depends on the text.[1] |  |
| The six accents listed to the right are the ones that could need stress help. These accents appear only at the start or end of their words. But a word with one of them at its start may not be initially stressed. And, likewise, a word with one of them at its end may not be finally stressed. So these accents need a helper to mark the stress in some words. | *segol*<br>*pashta*<br>*zarqa / tsinnor*<br>*telisha qetanah*<br>*telisha gedolah*<br>*deḥi* |

An accent and its helper typically match in shape. But they often differ in horizontal placement relative to their letter. To implement this, fonts need to use context to distinguish an accent from its helper. Fonts need to use context because except for *zarqa*,[2] an accent and its helper always share a code point. These distinctions are fragile. **This proposal will allow fonts to robustly distinguish an accent from its helper.**

[1] For detailed counts of possible helpers, by accent, see Section 9M.

[2] Regarding the use of *tsinnor* as a helper for *zarqa*, see Section 9G.

| | |
|---|---|
| Contextual distinctions are fragile because they can be defeated by mid-word changes to font weight, size, or color. For example, in the words shown to the right, the *segol* helper's centering was defeated by changes to the styling of the *tav* (ת) grapheme cluster, making it heavier, larger, and red, respectively.[3] |  |

A typical Biblical Hebrew font makes many contextual distinctions. So, distinguishing an accent from its helper is just one of many sources of fragility. But it is one of the most important sources of fragility. And, unlike many other sources of fragility, it stems from the need to make typographic distinctions that are motivated by semantic concerns.[4]

Unicode has code points for the **five final forms** of the Hebrew letters that have such forms. This proposal extends that design by adding code points for the six medial forms[5] of the accents that have such forms. (A helper for an accent *x* can be considered the medial form of *x*.)

Adding helpers would bring Unicode and the **Michigan-Claremont Encoding** (MCE) into closer sync.[6] Before Unicode, the MCE was the most important encoding of Biblical Hebrew, and it is still widely used. It is surprising that Unicode's Hebrew accent support was based just on SI 1489 and not also on the MCE.[7]

[3] Contextual distinctions can survive a color change in some software but not all. For example, they do not survive in Google Docs. Weight and size changes are lethal in all software I have tried.

[4] Many other sources of fragility come from the need to make typographic distinctions motivated by purely typographic concerns such as collision avoidance. Collision avoidance is of course a not a semantic concern, and as such would not typically be addressed by adding code points.

[5] Usually not literally a different form (shape); rather, a different horizontal placement.

[6] Regarding helpers in the MCE, see Section 9E.

[7] SI 1489 is from SII (מת"י), the Standards Institution of Israel (מכון התקנים הישראלי).

# 1. Background

In the vast majority of words in the Hebrew Bible, the stress is clear from the primary accent alone, because it is **an accent we can rely on** to appear on the "FLSS": the first letter of the stressed syllable.

| | |
|---|---|
| For example, in a word whose only accent is *munaḥ*, we can rely on it to appear on the FLSS, as shown to the right in Gen. 1:1 ברא. |  |
| A more complex example is a word with *munaḥ* and then *zaqef*. There, we know that *zaqef* is the primary accent and we can rely on it to appear on the FLSS,[8] as shown to the right in Gen. 1:14 ולמועדים. |  |

In a small minority of words in the Hebrew Bible, the stress is not clear from the primary accent alone, because it is **an accent we cannot rely on** to appear on (any letter of) the stressed syllable. I.e., it is one of the six accents that could need stress help.

- Such an accent may appear on the stressed syllable. It may even appear on the FLSS.
- But such an accent may appear off the stressed syllable.

To make the stress clear in a word with such an accent *x*, either:

- The word must have a helper for *x*.
- Or the word must appear in a context where helpers for *x* are used consistently, i.e., used whenever needed. **In such a context, an "un-helped" *x* is an accent we can rely on** to appear on the stressed syllable.

| | |
|---|---|
| For example, consider a word whose only accent is *telisha gedolah*, in a context where *telisha gedolah* helpers are not used consistently. The stress of such a word is unclear,[9] as shown to the right in three versions of Gen. 1:30 הארץ. Each version shows, in orange or green, a syllable that might have the stress. It happens to be the green one, but we have no way of knowing that without a helper. |  |

[8] We will not get into how we know that *zaqef* is this word's primary accent.

[9] In some cases the stress may be clear, but not from the word's accent alone.

| | |
|---|---|
| Now consider Gen. 1:12 דשא, shown in two versions to the right. Its stress is unclear unless we know we are in a context where *telisha gedolah* helpers are used consistently. |  |

These problems also apply to compound words. For example, the stress can be unclear in a compound word whose only accent is a *telisha gedolah*, as in the two words we explore below.

| | |
|---|---|
| Gen. 6:19 ומכל־החי is shown in two versions to the right. Its stress is unclear without a helper. |  |
| Josh. 10:37 ויכוה־לפי־חרב is shown in two versions to the right. Its stress is unclear unless we know we are in a context where *telisha gedolah* helpers are used consistently. |  |

As mentioned above, all six accents that could need stress help appear only at the start or end of their word (or word-part).[10] An accent that appears only at the start of its word (or word-part) is **prepositive**. An accent that appears only at the end of its word is **postpositive**. We now further define "prepositive" and "postpositive."

On a non-compound word, a **prepositive** accent is placed on (or slightly before) the first letter. On a compound word, it is placed on (or slightly before) the first letter of the stressed part of the compound. (The stressed part of the compound may or may not be the first part of the compound.) There are above-prepositives and below-prepositives. Even in words with initial stress, the placement of prepositives is often unlike the placement of normal accents.[11] It is often better to think of a prepositive as applying to its word as a whole even though in terms of Unicode and fonts, it applies only to the first letter of its word.

A **postpositive** accent is placed on (or slightly after) the last letter of its word. This is true whether the word is compound or not. There are only above-postpositives, no below-postpositives. Even in words with final stress, the placement of postpositives is often unlike the placement of normal

[10] Some accents appear only at the start but never need stress help. See Section 9B.

[11] Regarding the special placement of prepositives, see Section 9H.

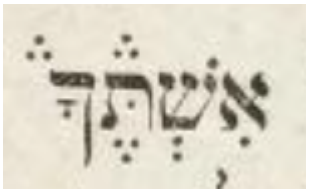
accents. It is often better to think of a postpositive as applying to its word as a whole even though in terms of Unicode and fonts, it applies only to the last letter of its word.

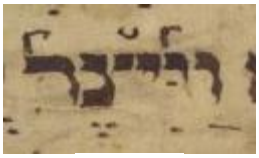
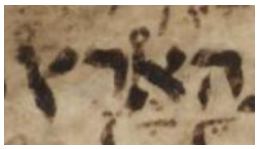
The helper for an accent *x* is typically shaped just like *x*, but unlike *x*, it always appears on the FLSS. Also, *x* and its helper may differ in horizontal placement relative to their letter. In particular, the helper for *x* may be horizontally centered on its letter even when *x* itself is right-biased or left-biased relative to the first or last letter of its word, respectively. Prepositives are almost always right-biased, sometimes even appearing slightly before their letter. Postpositives are often (but far from always) left-biased, sometimes even appearing slightly after their letter.

Let us substitute in *pashta* for the variable *x* above, to clarify these abstract statements with a concrete example. The *pashta* helper is typically shaped just like *pashta*, but unlike *pashta*, it always appears on the FLSS. Also, the *pashta* helper may be horizontally centered even when *pashta* itself is left-biased.

| | |
|---|---|
| For example, helpers for *pashta* and *telisha gedolah* can be used in Gen. 1:7 המים and Gen. 1:30 הארץ, as shown to the right. Most or all editions have the *pashta* helper on המים; only some editions have the *telisha gedolah* helper on הארץ. |  |

**For accents other than *pashta*, the most authoritative Tiberian manuscripts use helpers infrequently or not at all.**

| | |
|---|---|
| In contrast, for the last two centuries or so, some reader-friendly editions have used helpers consistently, for all five relevant accents in the prose system. For example, to the right we return to Gen. 3:17 אשתך, showing that it has a *segol* helper in volume 1 of Heidenheim's חומש עין הסופר (*Ḥumash 'Ein ha-Sofer*), published in 1818. |  |
| To the right we show another example of Gen. 3:17 אשתך with a *segol* helper. This example is also old, but not as old as Heidenheim. It is from volume 1 of the Baer *Masoretic Bible*, published in 1869. |  |

| | |
|---|---|
| Some of the later, less-authoritative manuscripts in the Tiberian style used helpers for accents other than *pashta*, though not necessarily consistently. To the right we show *telisha qetanah* with its helper in Gen 2:7 וייצר in the Austrian National Library's Cod. Hebr. 4,[12] dated 1344. We also show this word typeset.[13] |  וַיִּיְצֶר |
| To the right we show another *telisha qetanah* with its helper, this time in Gen. 1:30 הארץ in the Vatican's Urb. ebr. 2, written around 1100 in Italy.[14] We also show this word typeset. |  הָאָרֶץ |

As mentioned above, when helpers are used consistently, an **absent helper** means that the word's stress is **characteristic**, by which we mean final stress for postpositives and initial stress for prepositives. When a word has characteristic stress, phonetics line up with visuals: phonetically, the stress is at the same extreme where the true accent is, visually.

| | |
|---|---|
| For example, we used Gen. 1:12 דשא as an example of unclear stress, but, as shown to the right, its initial stress is clear in the context of an edition in which *telisha gedolah* helpers are used consistently. | דֶּשֶׁא |

While many editions use helpers consistently for all five relevant accents of the prose system, few editions use helpers for the two relevant accents of the poetic system, *deḥi* and *tsinnor*. Nonetheless, Unicode should support poetic helpers. This support requires only a code point for the *deḥi* helper. Support for the *tsinnor* helper is "free" because the prose accent *zarqa* and the poetic accent *tsinnor* share a code point, U+05AE ZINOR (sic). Therefore the *zarqa* helper and the *tsinnor* helper can share a code point, too.

There is no consensus yet on the shape and placement of the *deḥi* helper.[15] Nonetheless, Unicode should support it, in the abstract, because there is emerging demand for it even in the absence of a consensus on its concrete details.

[12] It is also known as the Shlomo Halevi Vienna Bible.

[13] The way Cod. Hebr. 4 writes this word may be confusing. See Section 9D.

[14] Unusually, Urb. ebr. 2 has *telisha qetanah* rather than *telisha gedolah* here.

[15] Regarding the *deḥi* helper, see Section 9I.

# 2. Proposed characters

I propose the following new code points and annotations (H.A. = HEBREW ACCENT):

1. U+05XX+0 H.A. SEGOL STRESS HELPER
2. U+05XX+1 H.A. PASHTA STRESS HELPER
3. U+05XX+2 H.A. TSINOR STRESS HELPER
4. U+05XX+3 H.A. TELISHA QETANA STRESS HELPER
5. U+05XX+4 H.A. TELISHA GEDOLA STRESS HELPER
6. U+05XX+5 H.A. DEHI STRESS HELPER

U+05XX+0 H.A. SEGOL STRESS HELPER

• marks the stress in a word with the accent segol

→ U+0592 H.A. SEGOL

U+05XX+1 H.A. PASHTA STRESS HELPER

• marks the stress in a word with pashta

→ U+0599 H.A. PASHTA

U+05XX+2 H.A. TSINOR STRESS HELPER

• marks the stress in a word with 05AE ZINOR (sic), i.e., in a word with the prose accent zarqa or the poetic accent tsinor

→ U+05AE H.A. ZINOR

U+05XX+3 H.A. TELISHA QETANA STRESS HELPER

• marks the stress in a word with telisha qetana

→ U+05A9 H.A. TELISHA QETANA

U+05XX+4 H.A. TELISHA GEDOLA STRESS HELPER

• marks the stress in a word with telisha gedola

→ U+05A0 H.A. TELISHA GEDOLA

U+05XX+5 H.A. DEHI STRESS HELPER

• marks the stress in a word with dehi

→ U+05AD H.A. DEHI

The annotations I propose in “Re-documenting ZARQA and ZINOR” should be adjusted if this proposal is accepted, since those annotations prescribe “self-help” for ZINOR, i.e., they prescribe that ZINOR (and only ZINOR) should be used as the helper for ZINOR.

# 3. Properties

The properties of each helper code point would be the same as the properties of their true accent, except for TSINOR STRESS HELPER. It would have the generic “above” combining class (230) rather than the “above-left” combining class (228) of its true accent, U+05AE ZINOR. That combining class 228 is only used for ZINOR and, I feel, should not even have been used there. It certainly should not be used for TSINOR STRESS HELPER.

# 4. Collation

The collation of each helper code point would be the same as the collation of their true accent.

# 5. References

(None listed here: references are given either inline, or in footnotes, or as the destination of hyperlinks.)

# 6. Acknowledgments

Thanks to Seth (Avi) Kadish for his help with this proposal, particularly for providing references to Heidenheim and Baer.

# 7. Examples

| | |
|---|---|
| In the new Trope Trainer software,[16] *pashta* and its helper differ in shape, as shown to the right. It is rare for an accent and its helper to differ like this. Though rare, this is the kind of distinction that this proposal would make robust. |  |
| In the new Trope Trainer, *zarqa* and its helper differ in color.[17] It is rare for an accent and its helper to differ like this. Though rare, this is the kind of distinction that this proposal would make robust, if implemented in a color font.[18] |  |
| In the *Tiqqun* of Feldheim's "Simanim" imprint, some letters are larger than others.[19] Some of these larger letters have helper accents, as shown to the right in the Gen. 3:17 word אשתך that we have often used as an example, and in the word that follows, ותאכל. Enlargement usually defeats a font's "smarts" that distinguish an accent from its helper, so such styling is a motivation for this proposal. |  |
| The *Tiqqun* uses left-biased placement for *segol*, but centers its helper. This can be seen in the אשתך example above, but is more clearly seen by comparison with Deut. 6:11 לא־מלאת, shown to the right, which shows the left-biased placement of true *segol* on *tav* (ת). (By chance, לא־מלאת also has a large letter with a helper accent, a large *lamed* (ל) with a *segol* helper.) |  |

As shown in Gen. 3:17 ותאכל above, the *Tiqqun* usually uses left-biased placement for both *pashta* and its helper, so a font trying to mimic the *Tiqqun* might not need to distinguish *pashta* from its helper, at least in this word. But in some words with *lamed*, the distinction might be needed.

[16] Not the original Trope Trainer software by Kinnor (Buchler).

[17] Admittedly, this may be an accident: see Section 9A.

[18] To put it more fully: this proposal would make the distinction robust compared to "ZINOR self-help." Though I oppose the use of ZARQA as the ZINOR helper, I have to admit that it does allow ZINOR to be robustly distinguished from its helper.

[19] These large letters are unrelated to the two dozen or so large letters that are officially part of the Hebrew Bible. See Section 9F.

| | |
|---|---|
| On *lamed*, the *Tiqqun* centers the *pashta* helper, presumably because of the ascender of the *lamed*. This is shown to the right in Deut. 6:23 לנו. I have not found an example in the *Tiqqun* where *lamed* with a *pashta* helper is large, but there is no reason it could not exist. |  |
| In some editions, Esth. 9:9 פרמשתא has a *mem* (מ) with a *pashta* helper and then a small *shin* (ש), as shown to the right in a Koren Tanakh. |  |
| If a font has "smarts" that distinguish *pashta* from its helper, they will usually be defeated by this size change, unless the font handles this particular situation explicitly. As shown to the right, my font, Taamey D, handles this situation explicitly in the fully pointed case (the upper case). But in the accent-only case (the bottom case), my font's "smarts" are defeated: the *pashta* helper slips to the left, as if it were a true *pashta*. |  |

# 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. Supplements

## 9A. Trope Trainer and *zarqa*

Trope Trainer colors *zarqa* and its helper differently. This may be an accident, since Trope Trainer is probably based on a text that uses different code points for *zarqa* and its helper. Unaware of this quirk in its base text, and unaware of quirks in Unicode naming, Trope Trainer's developers may, quite sensibly (but incorrectly) have written code to color only U+0598 ZARQA and not also U+05AE ZINOR. (Judging from its name and its location in the Hebrew block, ZINOR seems like it would be a poetic-only accent, and therefore irrelevant to Trope Trainer.) So ZINOR may be accidentally left black rather than intentionally colored black. This may show the continuing confusion surrounding these code points, related not only to mistakes in their names but also to annotations that seem to only add to the confusion. Hence my proposal, "Re-documenting ZARQA and ZINOR."

## 9B. Two prepositive accents never need stress help

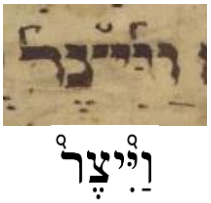
Although all accents that could need stress help are prepositive or postpositive, the converse is not true: the prepositives *yetiv* and *geresh muqdam* never need stress help.

*Yetiv* never needs stress help because it only appears on words with initial stress. Therefore *yetiv* always clearly (though trivially) marks the stress.

*Geresh muqdam* never really needs stress help because it is always paired with a *revia* that marks the stress. At least, *geresh muqdam* is always paired with a *revia* in theory. In practice, when stress is initial, often just the *geresh muqdam* appears. This simplified notation still marks the stress. In fact, if this simplified notation is used consistently, the *revia* of *revia mugrash* functions like a helper for *geresh muqdam*!

## 9C. [removed]

# 9D. Gen 2:7 וייצר in Cod. Hebr. 4

| | |
|---|---|
| This manuscript's way of writing וייצר may be confusing. |  וַיִּיצֶר |
| There is a Masorah circle above and centered on this word, but it is shaped more like a "Masorah cup." (The circle is incomplete: only its bottom half or so appears.) |  |
| Also confusing is the letter *tsadi*, which is broken. Or perhaps it is not quite broken, but its two parts are connected only by a very thin strand of ink. The result looks like the top of what might be a third *yod* (י) followed by a shape like a *nun* (נ). |  |
| The *telisha qetanah* accent and its helper do not have the circular shape used in the most authoritative manuscripts. Nor do they have the "circle on a stem" (lollipop or looking glass) shape used in most printed editions. |  |

## 9E. Helpers in the Michigan-Claremont encoding

Before describing helpers in the MCE, I will briefly describe the goal of the MCE. That goal is far more constrained than Unicode's goal.[20] The goal of the MCE is to provide an encoding suitable for an important digital edition of the Hebrew Bible, the Westminster Leningrad Codex (WLC).[21] So, to describe the goal of the MCE we must really describe the original goal of the WLC.[22] The original goal of the WLC was to represent the body text of BHS, an important edition whose body text strives to transcribe the Leningrad Codex (LC).

| | |
|---|---|
| The MCE has dedicated codes for three of the six helpers we propose for Unicode. They are listed to the right. The 1982 document known as MCMOT[23] defines accents 24 and 44. The 1989 document known as the "Supplement to the MCMOT"[24] defines accent 33.[25] | 24: *tel. qet.* helper<br>44: *tel. ged.* helper<br>33: *pashta* helper |

Here is how "MCMOT" defines Accents 24 and 44:[26]

> *Teliša qaton*, usually postpositive (04), has a distinct code (24) when it is internal to a word.

[20] Indeed it is hard to describe Unicode as having a single goal. It would be easier to describe Unicode as having multiple goals, some of which are sometimes even in tension with each other!

[21] This edition is now known as the WLC, but it has been known by other names in the past.

[22] In recent years, the goal of the WLC shifted slightly. Its goal is now to represent the LC itself, even if this means diverging from BHS. I.e., the WLC now diverges from BHS in places where it views BHS as being in error. In practice this slight shift has not changed the needs of the encoding.

[23] "MCMOT": Code Manual for the Michigan Old Testament (Research Memorandum UM82-1). 19 March 1982. (Oddly, CMMOT would be the expected initialism for the title.) It appears as a text file named "michigan.man" in distributions of the WLC.

[24] "Supplement to the MCMOT": Supplement to the Code Manual for the Michigan Old Testament. The J. Alan Groves Center for Advanced Biblical Research (formerly known as the Westminster Hebrew Institute). Westminster Theological Seminary. Last Revised 6/7/89. It appears as a text file named "supplmt.wts" in distributions of the WLC.

[25] For all I know, Accent 33 may have appeared in the Supplement years before the Supplement's last revision in 1989. When it characterizes its "last revision" as being in 1989 (6/7/89), the supplement means "last revision with untracked changes"; the Supplement currently has tracked changes as recent as 2009.

[26] I have placed transliterations in italics and I have added a caron above the letter 's' whenever it appears in *telisa*, making *teliša*.

> *Teliša gadol*, usually prepositive (14), has a distinct code (44) when it is internal to a word.
>
> There are [a few] pairs of accents whose members have the same graphic form but which we have coded separately. These [include] two forms each of *teliša qaton* (04,24) and *teliša gadol* (14,44).

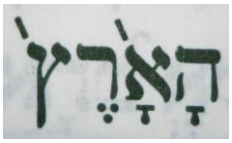
Here is how "Supplement to the MCMOT" defines Accent 33:[27]

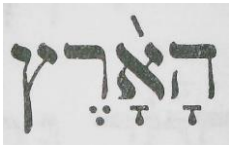
> Originally, [in tandem *pašta*,] the first mark was encoded as an *'azla* ('63' in the Michigan format) and the second as a *pašta* ('03' in the Michigan format). We have made the first mark a '33'. Note that it is written after the letter not over it (unlike the *'azla*).[28] The coding of the *pašta* remains '03'.

So, the addition of the dedicated *pashta* helper, Accent 33, was the result of a disunification of Accent 63. After disunification, Accent 63 became dedicated to *qadma*. The Supplement's note about *azla* (*qadma*) being "over" (above-center) its letter suggests that Accent 33 was **not** added because of a desire for a dedicated *pashta* helper per se. Rather, the quote suggests that Accent 33 was added because a dedicated *pashta* helper was the only way to meet the following goals:

- Semantically, they wanted *pashta* and its helper to be distinct. This goal was already met by sharing Accent 63 between *qadma* and the *pashta* helper.
- Graphically, they wanted the *pashta* helper to have an encoding prescribing above-left placement. This goal could **not** be satisfied by sharing Accent 63 because Accent 63 had to prescribe above-center placement to represent *qadma*.

The BHS images to the right show examples of an above-left *pashta* helper and an above-center *qadma* on the *alef* (א) of Gen. 1:11 and 50:11 הארץ.

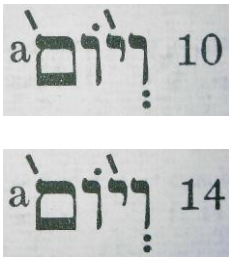




[27] I have placed transliterations in italics and I have added a caron above the letter 's' whenever it appears in *pasta*, making *pašta*.

[28] I would rephrase this as "Note that in BHS, the *pashta* helper is placed above-left, not above-center like *qadma*."

Even if BHS placed the *pashta* helper above-center as it does *qadma*, the MCE might have needed a dedicated *pashta* helper for a reason not mentioned in the Supplement:

| | |
|---|---|
| The reason is that in two words in BHS, a *qadma* appears before a *pashta*. These words are in the Decalogues, in Ex. 20:10 and Deut. 5:14.[29] They are shown to the right. This combination of *qadma* with *pashta* is normally impossible but happens here because BHS represents the Decalogues in the confusing dually-accented style, as in the most authoritative Tiberian manuscripts such as the LC. |  |

For these two words, a dedicated *qadma* is needed to avoid ambiguity between *qadma* and the *pashta* helper. In all other words, this ambiguity can be resolved word-locally, by a simple rule: when a "PH/Q" (a code shared by *pashta* helper and *qadma*) appears before a *pashta*, that PH/Q represents a PH rather than a Q. In these two words, the ambiguity can be resolved, but only by a rule that is too complex to be practical.[30] To keep things simple, a dedicated *qadma* is needed.

These two special words were encoded wrong, as 33…03 (*pashta* helper … *pashta*), until quite recently. (They were fixed in WLC 4.22). So these words did **not** motivate the addition of Accent 33. Indeed they probably came to be incorrectly coded as 33…03 in a global conversion of 63…03 to 33…03. But eventually these words took advantage of the addition of Accent 33, when they were fixed to be encoded back as they originally were, as 63…03!

From what we know of the MCE's *pashta* helper, it is clear why the MCE does **not** need a dedicated *zarqa* helper. The *zarqa* helper and *tsinnorit* can share Accent 82 because:

- In BHS, both the *zarqa* helper and *tsinnorit* are placed above-center.
- In BHS, there is no ambiguity between the *zarqa* helper and *tsinnorit*. E.g., there are no words analogous to those two odd Decalogue words.

In the Unicode context, I advise that this practice be deprecated, i.e. I advise that the code point for *tsinnorit* be dedicated to *tsinnorit*. Indeed, this is the main point of my proposal called "Re-documenting ZARQA and ZINOR." In contrast, in the MCE context, this practice seems fine to me since the MCE is, unlike Unicode, effectively a font-specific encoding, where it is okay to thoroughly conflate graphics and semantics.

[29] In other editions, the verse numbers are one less (Ex. 20:9 and Deut. 5:13).

[30] The rule could determine this PH/Q to be a Q because a PH in such a position would merely be re-iterating that the stress is final. (Indeed, because the stress is never on a vocal *shewa*, this is effectively a one-syllable word.) While there were some manuscripts that did this kind of re-iteration, the LC is not one of them, and therefore BHS does not do this.

The MCE does not need an encoding, either dedicated or shared, for helpers for *segol*, *tsinnor*, or *deḥi* because these helpers do not appear in BHS. (Presumably, they do not appear in the LC either).

| | |
|---|---|
| In summary, the MCE does not have dedicated codes for three of the six helpers we propose for Unicode. They are listed to the right. The MCE has a non-dedicated code for the *zarqa* helper, and has no code of any type for the rest. | *segol*<br>*zarqa / tsinnor*<br>*deḥi* |

## 9F. Officially large letters in the Simanim *Tiqqun*

| | |
|---|---|
| In the *Tiqqun*, these "officially large" letters appear in the non-pointed (scroll-like) (*ketiv*) column as opposed to our column of interest, the pointed (*qere*) column. For example, to the right we show two versions of Ex. 34:14 אחר, one from the non-pointed column, with a large *resh* (ר), and one from the pointed column, with a *resh* of normal size. |  |

## 9G. *Tsinnor* as the helper for *zarqa*

In most fonts, *pashta*, *zarqa/tsinnor*, and *deḥi* each have a centered “lookalike” accent. (Those lookalikes are *qadma*, *tsinnorit*, and *tarḥa*, respectively.) Thus, though I think it is a bad idea, the code points for those lookalikes can be used as helpers for *pashta*, *zarqa/tsinnor*, and *deḥi*. In practice, this is only done with *zarqa*: the code point for *tsinnorit* is widely used for the *zarqa* helper. In “Re-documenting ZARQA and ZINOR,” I propose that Unicode deprecate this practice. (Currently Unicode pretty much prescribes this practice.)

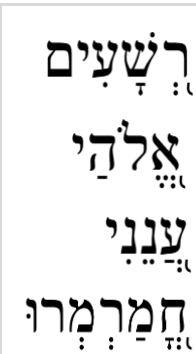
## 9H. The special placement of prepositives

Below we shall see how the special placement of prepositives supports the idea that though they appear at the start of their word, they "belong" only to their word as a whole, not necessarily to its first syllable. (This is analogous to the placement of postpositives, which supports the idea that though they appear at the end of their word, they "belong" only to their word as a whole, not necessarily to its last syllable.)

| | |
|---|---|
| Unlike all other below-accents, a below-prepositive is placed **before** any vowel it "contends" with. Examples of this abound, and appear as early in the poetic books as the *deḥi* in Ps. 1:1 חטאים, shown to the right.[31] |  |

As mentioned in Section 9I, this creates a big question about how a *deḥi* helper should be placed, relative to any vowel it might contend with. Contrast this with the minor stylistic choice of whether the helper for an above-accent should mimic its true accent's left- or right-biased placement. The helper for an above-accent must merely dodge the occasional *ḥolam* dot, *shin* dot, or *sin* dot.

Below we will give examples of prepositives appearing in places that we know cannot have the stress, i.e. cases where we know a helper is needed. Though needed, whether it is present is a separate question. In the examples we give below, the needed helpers are present for *telisha gedolah* and missing for *deḥi*, as is often the case.

| | | |
|---|---|---|
| If a word has an initial vocal *shewa*, the only primary accent that can appear at the start of that word is a prepositive.[32] Examples of this abound. To the immediate right we show four *deḥi* examples, from Ps. 1:5, 7:2, 13:4, and Job 16:16.[33] Further to the right we show four *telisha gedolah* examples, from Gen. 8:22, Deut. 9:6, Gen. 7:2, and 2 Sam. 24:13.[34] |  |  |

[31] We show no examples using *yetiv* because that accent is not relevant to our proposal. (*Yetiv* is the only below-prepositive other than *deḥi*.)

[32] Not even all prepositives can do this: *yetiv* cannot do this. (But regarding *yetiv* and **silent** *shewa*, see Section 9J.) Regarding the few cases of non-primary accents on an initial vocal *shewa*, see my external document, "*Tsinnorit & Oleh* on Initial Vocal *Shewa*."

[33] The Job word happens to be part of a *ketiv/qere*.

[34] We show no examples using *geresh muqdam* because that accent is not relevant to our proposal. (*Geresh muqdam* would complete the list of the three prepositives other than *yetiv*.)

| | | |
|---|---|---|
| Relatedly, if a word has an initial vocal *shuruq*, the only primary accent that can appear at the start of that word is a prepositive.[35] To the right we show examples from Ps. 22:16 and Gen. 7:7. | וְלִשׁוֹנִי | וּבָנָיו |

Here are some other examples of prepositives appearing in places that we know cannot have the stress:

- If a word has an initial *qamats qatan*, the only primary accent that can appear at the start of that word is a prepositive.[36]
- If a word has an initial *gaʿya* (*meteg*), the only primary accent that can appear at the start of that word is a prepositive.[37]

Thus, even though in terms of Unicode and fonts, prepositives "belong" to the first letter of their words, their special placement supports the idea that they "belong" only to their word as a whole.

[35] This is another case where not even all prepositives can do this: this is another case where *yetiv* cannot do this. Regarding the few cases of non-primary accents on an initial *shuruq*, see Section 9K.

[36] Two points:
- Regarding non-primary accents with *qamats qatan*, see Section 9L.
- There are two possible exceptions to this rule: see Section 9L.

[37] In a few cases, the non-primary accent *oleh* appears with a *gaʿya*.

## 9I. The *deḥi* helper

It is not clear how to best represent the *deḥi* helper. One challenge is that *deḥi* is the only below-prepositive that could need stress help,[38] so there is no precedent for whether its helper should appear before or after a vowel it "contends" with.

Another challenge is that *deḥi* and *tarḥa* almost always share a shape. So, though accents and their helpers almost always share a shape, it may cause confusion if *deḥi* and its helper share a shape, because of the chance of confusing the *deḥi* helper with *tarḥa*. There is an analogous chance of confusing the helpers for *pashta* and *tsinnor* with other accents,[39] but the below-prepositive nature of *deḥi* makes things worse. It makes things worse since if the *deḥi* helper appears before any vowel it "contends" with, it can look like a *tarḥa* on the previous letter. In such cases, we might not only confuse the accent but also its location!

[38] The only other below-prepositive is *yetiv*, and it never needs stress help.

[39] The *pashta* helper can be confused with *qadma* and the *tsinnor* helper can be confused with *tsinnorit*. (The *zarqa* helper cannot be confused with *tsinnorit* since they come from different systems: the prose and poetic systems, respectively.) The chance of confusing the *pashta* helper with *qadma* is well known and well-educated-against. The chance of confusing the *tsinnor* helper with *tsinnorit* is not well known since few editions use a *tsinnor* helper.

## 9J. *Yetiv* on an initial silent (!) *shewa*

| | |
|---|---|
| Four instances of the word שתים have *yetiv*, as shown to the right.[40] Extraordinarily, this word's initial *shewa* is usually considered to be silent. (This is true of all instances of שתים, not just these four with *yetiv*.) (It is the only word in which an initial *shewa* is usually considered to be silent.) Thus, while *yetiv* never appears at the start of a word with an initial **vocal** *shewa*, *yetiv* does appear at the start of four instances of a word with an initial **silent** *shewa*! |  |

[40] The four instances of שתים with *yetiv* are Lev 23:17, Ez 1:11, and Ez 41:24 (2×).

## 9K. Non-primary accents on initial *shuruq*

In some manuscripts and printed editions, *geresh* appears on initial *shuruq* in Ezek. 48:10 וּלְאֵלֶּה. Whether *geresh* is this word's primary accent is unclear. If *geresh* is this word's primary accent, this would be an exception to my claim that the only primary accents that can appear "on" an initial *shuruq* are prepositives. (This word's accentuation varies, but it always involves at least one *telisha gedolah* and at least one *geresh*. It is one of only five words that combine *telisha gedolah* with either *geresh* or *gershayim*.)

Secondary accents can appear on *shuruq*, but are quite rare: I find only 14 cases. Twelve of these 14 are *metigah* on *shuruq*. (Here as always, *metigah* appears only as part of the pair *metigah-zaqef*.) The remaining two of these 14 are *merkha* preceding the poetic accent *azla legarmeh*.

## 9L. Accents with *qamats qatan*

It is not rare for a non-primary accent to appear with *qamats qatan*. It is not rare for this to happen with *metigah* (57 cases), and though it is rare, it does happen with the following:

- *oleh* (4 cases)
- *merkha* (either 1 or 3 cases)
- *munaḥ* (4 cases)

(When it happens with *merkha* or *munaḥ*, they are functioning as secondary accents, except in editions with the two exceptions described below.)

In the body of this document, we stated that if a word has an initial *qamats qatan*, the only primary accent that can appear at the start of that word is a prepositive. There are two possible exceptions to this. In Ps. 35:10 and Pr. 19:7, in some editions, the word כל appears independently rather than as the first part of a compound. In editions where כל is independent, its *merkha* functions as the primary (and indeed only) accent. Therefore, in editions where כל is independent, the (primary) accent of כל appears with *qamats qatan*, yet the (primary) accent of כל is not prepositive.

# 9M. Counts of possible helpers, by accent

| | count | prepos.? | poetic? |
|---|---|---|---|
| *segol* | 170 | | |
| *pashta* | 3,827 | | |
| ZT *zarqa* [41] | 180 | | |
| ZT *tsinnor* [42] | 55 | | poetic |
| *tel. qet.* | 333 | | |
| *tel. gad.* | 1,000 | prepos. | |
| *deḥi* | 2,304 | prepos. | poetic |

(End of document.)

[41] *zarqa/tsinnor* used as *zarqa*

[42] *zarqa/tsinnor* used as *tsinnor*