Toward a Prosodic-Phrasing Algorithm for Tiberian Hebrew
Vincent DeCaen
draft 9 (June 2011)¹

Abstract

Algorithm

$\textbf{BINMAX} \gg \textbf{PAUSE} \gg \textbf{WRAP}, \textbf{LONGWD} \gg \textbf{ALIGN} \gg \textbf{BINMIN} \gg \textbf{SYNTAX}$
0. Introduction: Generative Masoretics

0.0. The remarkable isomorphy of Biblical Hebrew (BH) syntactic structure (DeCaen 1995) and the syntax of the Tiberian logogenic chant (Price 1990, 1996) cries out for a formal explanation. The seminal study by Dresher (1994) explains this striking isomorphy by means of an abstract, intermediate prosodic representation\textsuperscript{2} negotiated at the syntax-phonology interface.

0.1. Dresher thereby convincingly explains (a) how and why the chant deviates from BH syntax; (b) the otherwise bizarre function of the accents in directly demarcating the domains in which Tiberian Hebrew (TH) phonological rules operate; and (c) the compression and expansion phenomena of the accent systems in terms of the tempo and grain of a natural, spoken language.

0.2. Dresher’s prosodic theory has tremendous explanatory power \textit{pace} Churchyard (1999). The scope of the theory can be extended and expanded into an ambitious programme in generative metrics which I call \textit{Generative Masoretics} (figure 1).

0.3. From the many phonological and musical transformations can be deduced a detailed theory of TH metrical phonology (DeCaen 2008): a further interface intervening between the syntax and the music. The resulting metrical grids can be leveraged into a metrical theory of BH accentual-syllabic poetry (DeCaen 2009, 2011).

0.4. The implied complex boundary tones and down-stepping of highs and lows of the prosodic representation naturally project intonational contours: the inherent music of the spoken language. The parade of alternating highs and lows of these intonational contours is undoubtedly the source of the logogenic chant (DeCaen 1999, see also Weil 1995).
2. Generative Masoretics and Computation

2.0. The prospect of robust, detailed theories of TH prosodic and metrical representation opens out on a computational-linguistic vista. From arbitrary syntactic³ inputs, it should be possible to algorithmically project a prosodic representation. From the prosodic structure, it should be possible to algorithmically generate metrical grids. And from metrical grids it should be possible to algorithmically project the musical grids of the two accent systems via musical transformations.

2.1. The heart of such a computational project is the demarcation of phonological phrases (φ), aligning the edges of φ with syntactic edges based on ranked preferences and constraints. The optimality-theoretic (OT),⁴ node-correspondence-based Align/Wrap Theory of Selkirk (2000, 2005) supplies the required constraint-based and edge-based mapping.
3. Prosodic Phrasing of Simplified Inputs

3.0. It is not reasonable to try to draft a complex OT prosodic-phrasing algorithm for long verses at one go. A reasonable strategy is to simplify inputs as much as possible to identify the rudiments of such an algorithm.

3.1. One considerable simplification is to restrict inputs to intonational phrases and not whole verses. It appears that aligning intonational phrases trumps aligning phonological phrases, as would otherwise be expected. The nature and distribution of intonational phrases, as a first approximation, is given in DeCaen (2011b).

3.2. Another simplification is to restrict inputs to the short, regular intonational phrases of the simplest type of BH poetry. The short, end-stopped lines of Psalms 111-112 (DeCaen 2009, 2011a) and Job 3 (approximately 100 lines) offer a good place to start.

4. Continuous Dichotomy in OT Perspective

4.0. The essential feature that unites BH syntax and the syntax of the accent systems is *continuous dichotomy* or top-down binary branching. Such continuous dichotomy can be imposed by an absolute ranking of binary constraints (1) as in (2).

\[(1) \quad \text{BINARY MAXIMUM} = \text{BINMAX} \]
\[
\text{Any prosodic phrase at any level of the prosodic hierarchy is maximally binary-branching.}
\]
\[
\text{BINARY MINIMUM} = \text{BINMIN} \]
\[
\text{Any prosodic phrase at any level of the prosodic hierarchy is minimally binary-branching.}
\]

\[(2) \quad \text{BINMAX} \gg \text{BINMIN} \]
4.1. The absolute ranking of BinMAX ensures that no node in the prosodic representation is ternary-branching. Notice that a phonological phrase ($\varphi$) consisting of one phonological word ($\omega$) incurs a penalty under BinMIN. Violations of this constraint should only arise, all things being equal, where the input is odd-numbered.

4.2. N.B. Recursive phrasing is permitted to achieve conformity with BinMAX.

5. Two-Word Problem and “Long Words”

5.0. Continuous dichotomy imposed by (2) entails that two phonological words ($\omega$) will necessarily be phrased together as one phrase ($\varphi$). However, such is not always the case: prosodic length can trump BinMIN. Consequently, a “long word” (3) must intervene as in (4).

(3) **PROSODIC “LONG WORD” = LONGWD**
A “long” phonological word ($\omega$) must be aligned with its own phonological phrase ($\varphi$).

(4) **BINMAX >> LONGWD >> BINMIN**

5.1. N.B. The detailed definition of “metrical long word” or *metrical dipod* (DeCaen 2008), sufficient to explain the Tiberian musical transformations—or rather, lack thereof—is NOT necessarily the same as “prosodic long word” that is revealed here in the prosodic-phrasing algorithm for the poetic system of accentuation. More anon.

5.2. Job 3:21b is parsed first without the LONGWD constraint in (5), and then with the intervening LONGWD constraint in (6). (As will become clear below, both words are “prosodic long words”.)

(5) Job 3:21b *wáyyahpərúh mimmaṭmónim*

<table>
<thead>
<tr>
<th>$\omega \omega$</th>
<th>BINMAX</th>
<th>BINMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Φ (\omega \omega)</td>
<td></td>
<td><em>!</em></td>
</tr>
<tr>
<td>(((\omega) (\omega))</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(6) Job 3:21b wáyyahparūhû mimmaṭmônîm

<table>
<thead>
<tr>
<th></th>
<th>BINMAX</th>
<th>LONGWD</th>
<th>BINMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ω ω</td>
<td><em>!</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ω ω)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;ω ((ω) (ω))&gt;</td>
<td></td>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>

5.3. The other instance of a two-word phrase in the limited database is Job 3:22a. However, this token is rendered ambiguous by the obscuring musical transformation that is triggered by metrical monopod 'ēlē—gîl. (Nevertheless, as shown below, haśśǝmēḥîm is not a “prosodic long word” according to the relevant distinction in §7.2.)

6. Three-Word Problem and Syntactic Edges

6.0. Continuous dichotomy does not resolve the three-word problem in (7), setting to one side the odd long word. Because of the odd-numbered input, there is necessarily some violation of BINMIN.

(7) Three-Word Problem

<table>
<thead>
<tr>
<th>ω ω ω</th>
<th>BINMAX</th>
<th>BINMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ω ω ω)</td>
<td><em>!</em></td>
<td></td>
</tr>
<tr>
<td>((ω ω) (ω))</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>((ω) (ω ω))</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>((ω) (ω) (ω))</td>
<td><em>!</em></td>
<td>???</td>
</tr>
</tbody>
</table>

6.1. In such cases, the principal observation is that the prosodic phrasing is sensitive to morphosyntactic “right edges”: ALIGNXP (8). Representative examples are provided in (9) and (10).

(8) ALIGN R (XP, φ) = ALIGNXP
Align the right edge of an XP in syntactic (PF) representation with the right edge of a phonological phrase (φ) in phonological (PR) representation.
6.2. However, in rare cases, it appears that the verb V is enforcing a syntactic edge against expectations. Inspection shows that the simple application of ALIGNXP may be trumped by higher-order syntactic considerations. The edge of a clause (CP) imposes a greater degree of disjunction than the typical syntactic edge, corresponding in the isomorphic prosodic representation to the right edge of an intonational phrase (I): ALIGNCP (11). Conjunction of constituents must also be observed: CONJ (12). Representative cases are offered respectively in (13)⁵ and (14)⁶.

(11) ALIGN R (CP, IP) = ALIGNCP
    Align the right edge of any clause CP in syntactic (PF) representation with the right edge of some intonational phrase IP in phonological (PR) representation.

(12) CONJUNCTION OF XPS = CONJ
    The conjunction of XPs must be respected.
    Conjunction is left-recursive (N.B. Ps 112:4b).
6.3. Constraints such as (11) and (12) will multiply rapidly. The significant generalization that is being missed is that constituents of the larger XP must be kept together. Following Selkirk, a “wrap” constraint is introduced in (15).

\[(15) \text{W} \text{R} \text{A} \text{P} \text{X} \text{P} = \text{W} \text{R} \text{A} \text{P}\]

A phrase (XP) in syntactic (PF) representation (specifically, its terminal string) must be contained within a prosodic phrase of the corresponding level (Selkirk 2005) in phonological representation.

6.4. It should further be noted that there are many false friends in the application of ALIGNXP: morphologically “free” versus “bound” nominal forms do not necessarily signal edges. To take a simple example, there is no right edge between a noun and its following XP complement (AP, e.g., Ps 111:10b; PP, e.g., Ps 112:2b): by rule, N’ (read “N bar”) does not impose a prosodic right edge (15).

\[(15)\]

```
NP
  N'
  |  XP
  |  N
```
6.5. Thus, apparent exceptions to ALIGNXP are analyzed so that there is no right edge, because there is no intervening XP. Two examples involving the complements of the participle (Prt), and one the complements of adjective (A), are given in (16).

(16) Ps 111:8a \[\text{[[[Prt]}|\text{Prt}]|\text{PP}]|\text{Prt}]|\text{PP}\]
    Job 3:15b \[\text{[[[Prt]}|\text{Prt}]|\text{NP}]|\text{Prt}]|\text{NP}\]
    Job 3:19b \[\text{[N]}|\text{NP}]|\text{[[A]}|\text{A}]|\text{PP}]|\text{AP}\]

6.6. There are rare cases where ALIGNXP is unable to decide among candidates (18). In such cases, the final arbiter (default) is the Syntax (17).

(17) Syntax
    Prosodic phrasing is isomorphic with the syntactic structure.

(18) Job 3:18a \[\text{[[Adv}]|\text{Adv}]|\text{[[N]}|\text{NP}]|\text{V}]|\text{TP}]|\text{CP}\]

<table>
<thead>
<tr>
<th></th>
<th>ALIGNXP</th>
<th>SYNTAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\omega \omega \omega))</td>
<td>(\ast)</td>
<td>(\ast!)</td>
</tr>
<tr>
<td>(\varnothing ((\omega) (\omega \omega)))</td>
<td>(\ast)</td>
<td></td>
</tr>
</tbody>
</table>

7. Prosodic “Long” Word Again

7.0. Ex hypothesi, the only way a line can escape the constraints outlined above is if a poetic “prosodic long word” is present. Only one three-word instance is found in the restricted corpus: the PP \(\lambda\nipl\o'\dot\text{t\i}yw\) is “long” ex hypothesi in Ps 111:4a (19).
In this particular case, there is an **exact minimum pair**: NP V PP in Ps 111:5a. Even more interesting, the PP in the latter case *lîrē’āyw* is also a TH “metrical long word” (metrical dipod). In other words, from a TH metrical point of view, the two PPs are metrically identical (20):

(20)  
\[
\begin{array}{cccc}
  & x & x \\
(x & x) & (x & x) & (x & x) & (x & x) \\
(l̄) & nip & ɾ̄ & t̄āyw & r̄ & ‘āyw
\end{array}
\]

However, the PPs are by no means identical from an **accentual-syllabic** point of view. The accentual-syllabic parse employed in DeCaen (2011a) is projected downwards in (21).

The “prosodic long word” *ləniplə ‘ōtāyw* has a **complete group or foot**; whereas the defective word *lîrē’āyw* has an **incomplete group** or degenerate foot. Thus, the ambiguous case in Job 3:22a (22), noted above at §5.3, is disambiguated: it is not a “prosodic long word”, despite being a “metrical long word”.

(21)  
\[
\begin{array}{cccc}
  & x & x & x \\
(x & x) & (x & x) & (x & x) \\
l̄ & nip & r̄ & t̄āyw & r̄ & ‘āyw
\end{array}
\]

---

(19) Ps 111:4a  **NP V PP**

<table>
<thead>
<tr>
<th></th>
<th>BinMax</th>
<th>LongWd</th>
<th>BinMin</th>
<th>AlignXP</th>
<th>Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ω ω  ω)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ω (ω  ω))</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>(ω (ω ω))</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7.3. The constraint LONGwd is accordingly modified as in (23).

(23) \textbf{PROSODIC “LONG WORD” = LONGwd}

A phonological word \((\omega)\) is \textit{“long”} if, with reference to its accentual-syllabic parse, its secondary foot is \textit{“complete”}.

A \textit{“long”} phonological word \((\omega)\) must be aligned with its own phonological phrase \((\varphi)\).

8. Four-Word Problem

8.0. Continuous dichotomy will always select—absent a “prosodic long word”—the candidate \(((\omega \omega) (\omega \omega))\) without some intervening constraint(s), as can be seen from the pruned tableau in (24). One way to think of BinMin is as a \textit{prosodic default} among a block of lowest-ranking defaults, including the proposed Syntax.

(24)

<table>
<thead>
<tr>
<th>(\omega \omega \omega \omega)</th>
<th>BinMax</th>
<th>BinMin</th>
</tr>
</thead>
<tbody>
<tr>
<td>(((\omega \omega) (\omega \omega)))</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>(((\omega \omega \omega) (\omega)))</td>
<td><em>!</em></td>
<td><em>!</em></td>
</tr>
<tr>
<td>(((\omega) (\omega \omega) (\omega)))</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>(((\omega) (\omega) (\omega \omega)))</td>
<td><em>!</em></td>
<td><em>!</em></td>
</tr>
<tr>
<td>((\omega \omega \omega \omega))</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>(((\omega) (\omega) (\omega) (\omega)))</td>
<td>*!</td>
<td>****</td>
</tr>
</tbody>
</table>

8.1. Sitting immediately above the “defaults” must be the basic alignment constraints.

Following DeCaen (19xx: paper on autolexical NP), a block of alignment constraints \texttt{ALIGN} (25)
is proposed. ALIGNXP (26) was already introduced above in (8). ALIGNX (27) has the effect of treating syntactic heads as *prosodic clitics*. The simple case of right recursion in (28) is consequently mirrored in the right-recursive nesting of phonological phrases in (29).

\[(25)\] \textbf{ALIGN} = ALIGNXP, ALIGNX

\[(26)\] ALIGN R (XP, φ) = ALIGNXP
Align the right edge of an XP in syntactic (PF) representation with the right edge of a phonological phrase (φ) in phonological (PR) representation.

\[(27)\] ALIGN L (X, φ) = ALIGNX
Align the head X of an XP in syntactic (PF) representation with the left edge of a phonological phrase (φ) in phonological (PR) representation.

\[(28)\]
```
NP
  N
  NP
  N
  NP
  N
  NP
  N
```

\[(29)\] \( [N [N [N [N]_NP]_NP]_NP]_NP \)

<table>
<thead>
<tr>
<th>(ω \ ω \ ω \ ω)</th>
<th>ALIGNX</th>
<th>BINMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>(((ω \ ω) (ω \ ω)))</td>
<td>*!</td>
<td>**</td>
</tr>
<tr>
<td>(((((ω \ ω) (ω)) (ω)))</td>
<td>*!</td>
<td>**</td>
</tr>
<tr>
<td>(((ω) ((ω \ ω)) (ω)))</td>
<td>*!</td>
<td>**</td>
</tr>
<tr>
<td>(((ω) ((ω \ ω) (ω))))</td>
<td>*!</td>
<td>**</td>
</tr>
<tr>
<td>(&amp;) (((ω)) ((ω) (ω ω)))</td>
<td>**</td>
<td>**</td>
</tr>
</tbody>
</table>
8.2. Examples involving ALIGNXP and ALIGNX are given in (30) and (31) respectively.

(30) Ps 112:4a V PP NP PP

<table>
<thead>
<tr>
<th></th>
<th>ALIGNXP</th>
<th>BinMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((ω ω) (ω ω))</td>
<td>!**</td>
</tr>
<tr>
<td>C</td>
<td>(((ω ω) (ω)) (ω))</td>
<td>!**</td>
</tr>
<tr>
<td></td>
<td>(((ω) (ω ω)) (ω))</td>
<td>!**</td>
</tr>
<tr>
<td></td>
<td>(((ω) (ω) (ω)))</td>
<td>!**</td>
</tr>
<tr>
<td>C</td>
<td>(((ω) (ω ω)))</td>
<td>!**</td>
</tr>
</tbody>
</table>

(31) Ps 112:8b P C V PP

<table>
<thead>
<tr>
<th></th>
<th>ALIGNX</th>
<th>BinMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((ω ω) (ω ω))</td>
<td>!**</td>
</tr>
<tr>
<td>C</td>
<td>(((ω ω) (ω)) (ω))</td>
<td>!**</td>
</tr>
<tr>
<td></td>
<td>(((ω) (ω ω)) (ω))</td>
<td>!**</td>
</tr>
<tr>
<td></td>
<td>(((ω) (ω) (ω)))</td>
<td>!**</td>
</tr>
<tr>
<td>C</td>
<td>(((ω) (ω ω)))</td>
<td>!**</td>
</tr>
</tbody>
</table>

8.3. The phrasing will ensure that arguments and clauses are not split over prosodic phrases, but rather there is some basic integrity. Ensuring top-down integrity falls to the block of “wrap” constraints in (15), where Wrap >> Align. As noted, this “wrap” function subsumes a myriad of preferences such as ALIGNCP and CONJ. An example with two XPs that must be “wrapped” is given in (32).

(32) Ps 111:7a [N NP]NP [NP [& NP]]&P

<table>
<thead>
<tr>
<th></th>
<th>WrapXP</th>
<th>ALIGN</th>
<th>BinMIN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(((ω ω) (ω ω)))</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>C</td>
<td>(((ω ω) (ω)) (ω))</td>
<td><em>!</em></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>(((ω) (ω ω)) (ω))</td>
<td><em>!</em></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>(((ω) (ω) (ω)))</td>
<td><em>!</em></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>(((ω) (ω ω)))</td>
<td><em>!</em></td>
<td>*</td>
</tr>
<tr>
<td>C</td>
<td>(((ω) (ω ω)))</td>
<td><em>!</em></td>
<td>*</td>
</tr>
</tbody>
</table>

8.4. The most important top-down constraint is TH “pause” or prosodic disjunction. TH pause is triggered by a number of syntactic structures subsumed under Comma Phrase (Selkirk 2005). As proposed in DeCaen (2011b), there are three ranked constraints (26)-(29) that can
taken as a block PAUSE (33), only dominated by BinMAX (1), but crucially adding here a default that aligns a Comma Phrase with a phonological phrase (37). The example of Job 3:3b is given in (38).

(33) \[ \text{PAUSE} = \text{ALIGNCOMMAP2} \gg \text{BINMIN}(I) \gg \text{ALIGNCOMMAP1} \gg \text{ALIGNCOMMAP}(\phi) \]

(34) \[ \text{ALIGN (COMMAP1, I)} = \text{ALIGNCOMMAP1} \]
Align the edges of a CommaP1 in syntactic (PF) representation with the edges of a in phonological (PR) representation.

(35) \[ \text{ALIGN (COMMAP2, I)} = \text{ALIGNCOMMAP2} \]
Align the edges of a CommaP2 in syntactic (PF) representation with the edges of a corresponding intonational phrase in phonological (PR) representation.

(36) \[ \text{BINMINIMUM (I)} = \text{BINMIN(I)} \]
An intonational phrase (I) must consist of at least two phonological phrases (\(\phi\)).

(37) \[ \text{ALIGN R (COMMA PHRASE, } \phi \text{)} = \text{ALIGNCOMMAP}(\phi) \]
Align the right edge of an XP in syntactic (PF) representation with the right edge of a phonological phrase (\(\phi\)) in phonological (PR) representation.

<table>
<thead>
<tr>
<th>(\omega \omega \omega \omega)</th>
<th>PAUSE</th>
<th>WRAP</th>
<th>ALIGN</th>
<th>BinMin</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\phi \ (\omega \ (\omega \ (\omega \ (\omega)))))</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>(\phi \ (\omega \ (\omega \ (\omega))))</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>(\phi \ (\omega \ (\omega \ (\omega))))</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
<tr>
<td>(\phi \ (\omega \ (\omega \ (\omega))))</td>
<td>*!</td>
<td>*</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

(38) Job 3:3b

<table>
<thead>
<tr>
<th>[NP</th>
<th>V_{amar}]CommaP</th>
<th>[V NP]CPCommaP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\omega \omega \omega \omega)</td>
<td>PAUSE</td>
<td>WRAP</td>
</tr>
<tr>
<td>(\phi \ (\omega \ (\omega \ (\omega))))</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>(\phi \ (\omega \ (\omega \ (\omega))))</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>(\phi \ (\omega \ (\omega \ (\omega))))</td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>

13
Bibliography


http://homes.chass.utoronto.ca/~decaen/papers/


End Notes

1 Acknowledgements...

2 To be clear, the accentuation is not itself a prosodic representation. This might be the mistaken impression from reading, e.g., Churchyard (1999: 222-252). The way the accentuation varies relative to the syntax simply suggests that it is sensitive to an abstract intermediate syntax that Dresher (1994) identifies as prosodic. We know independently that the accentuation has its own rhyme and reason, and has its own constraints and transformations. There are thus several intermediate “levels” separating the syntax from the music.

3 [make note of Kirk Lowery’s work on the syntax parser, update, prospects.]

4 On Optimality Theory, start with Prince & Smolensky (2004); then see Archangeli & Langendoen (1997) and Kager (1999).

5 See also Ps 112:9a, 10b; Job 3:11b, 21a.

6 See also 111:3a, 4b, 8b, 9c.