

The Word In Tiberian Hebrew*

Bezalel Elan Drescher

1. Introduction

The notion of a ‘word’ is subject to various ambiguities. The text of the Hebrew Bible (Elliger and Rudolph 1977) provides us with several conflicting notions of word. The Biblical text consists of two main layers, and the written word - in the sense of letters surrounded by blank space - differs in each layer.

The earlier layer contains a consonantal text, devoid of almost all indications of vowelling and punctuation. I will show that the word in this consonantal layer corresponds to a potential prosodic word, that is, a unit that could be an independent word for purposes of phrasing, whether or not it actually functions as such in any particular context. Since such words are not necessarily prosodic words in every context, I will call the word in this layer an orthographic word. Ordinary written Hebrew makes similar word divisions.

To more precisely indicate correct pronunciation and phrasing, various diacritic marks, or ‘points’, were later added to the consonantal text, producing a ‘pointed’ text. These marks include vowel signs, some allophonic consonantal distinctions, and an elaborate system of ‘accents’ that indicates position of stress, division into verses, and a highly articulated prosodic parse of each verse. Orthographic words of the consonantal text (potential prosodic words) could be joined together by hyphens to create a larger unit, the (actual) prosodic word. As we shall see, these constituents are prosodic words in the sense that they count as words for purposes of phrasing. I will show that the principles for forming

prosodic words – rules of cliticization – are quite complex, and interact in intricate ways with other aspects of prosodic structure, such as the phonological phrase and the intonational phrase.

Turning to the evidence of the phonology, I will distinguish between the prosodic word and the phonological word, which is the notion of word referred to by the phonology proper (segmental processes, syllabification, stress), as opposed to the phrasing. Though the phonological word necessarily has some relation to the prosodic word, the two concepts are not identical. Thus, a study of the word in Biblical Hebrew bears on issues of the syntax-phonology mapping in contemporary linguistic theory, as well as on the notion of levels in Lexical Phonology and Morphology.

2. The Consonantal Text: The Orthographic Word

In the consonantal text, all content words, such as nouns, verbs, adjectives, and numerals, are separate words, separated by a space from adjacent words. Most prepositions are also written as independent words. Prepositions that consist of only a single consonant (or consonant plus schwa, depending on whether the schwa is analyzed as inserted by rule or part of the underlying form), however, are written as bound prefixes, with no space separating them from what follows. It is clear that word status is not connected to semantics in this case, because all these prefixes have variants or synonyms consisting of more phonological material, and these are invariably written as independent words:

(1) Prepositions

<u>Prefix</u>	<u>Independent word</u>	<u>Gloss</u>
a. b(ə)-	bəmo: (poet.), ʕim	‘in, at, with, by’
b. l(ə)-	ləmɔ: (poet.), ʕel	‘to, for’
c. k(ə)-	kəmo: (poet.)	‘like, as’

Morphemes of the form $C(ə)$ do not make up a full syllable. Therefore, the above observations suggest the generalization that morphemes that consist of less than a full syllable are not written as independent words.

The prepositions in (1) have variants of the form $Ci-$ when prefixed to words that would otherwise have an initial syllable with a schwa:

(2) Variants of prefixed prepositions

a. <u>Cə-</u>	b. <u>Ci-</u>
bəða:vá:r	biðvar (unprefixed: dəvar)
‘in a word, matter’ (absolute)	‘in the matter of’ (construct)

The $ə \sim i$ alternation in these prepositions is phonological, with i being the default vowel in closed syllables.¹ The replacement of schwa by the full vowel i does not change the bound prefixal status of the prepositions.

In this category we can also place the conjunctive $w(ə)-$ ‘and’, called *waw* after the Hebrew letter used to represent it. This morpheme is written $u:-$ before a syllable containing schwa or a labial consonant, but remains a prefix:

(3) Conjunctive *waw*

a. <u>Elsewhere</u>	b. <u>Before schwa</u>	c. <u>Before labial</u>
wəða:wí:ð	u:xəná:ʕan	u:miryá:m
‘and David’	‘and Canaan’	‘and Miriam’

The *u:-* allomorph of the conjunctive forms an exception to an otherwise regular rule that Hebrew syllables, and hence words, begin with a consonant. Thus, *u:-* does not form a proper syllable, and so also falls under the generalization that morphemes that consist of less than a full syllable are not written as independent words.

Phonological subminimality is a sufficient condition for a morpheme to be written as an affix, but it is not a necessary condition. The preposition *min* ‘from, out of, than’ frequently occurs in the form *mi-*, where the final *n* has been historically assimilated to the following consonant. Synchronically, the allomorph *mi-* causes gemination of a following consonant and is always written as a prefix:

(4) The preposition *min*

a. <u><i>min</i></u>	b. <u><i>mi-</i></u>	c. <u><i>me:-</i></u>
min ha:ʕé:ʂ	middəváš	me:ʕé:ʂ
‘from the tree’	‘than honey’	‘from a tree’

Before gutturals, which do not normally geminate, the vowel is lengthened to *e:*. However, this lengthening does not restore the prefix to the status of independent word.²

Another morpheme with the shape CV that is always written as a prefix is the definite article *ha-*. Like *mi-*, *ha-* causes gemination of the following consonant; when gemination is not possible, the vowel of the article is lengthened to *a:*.³

(5) The definite article *ha-*

a. <u><i>ha-</i></u>	b. <u><i>ha:-</i></u>
hammélex ‘the king’	ha:ʕéveð ‘the servant’

The interrogative *ma* ‘what’ occurs in a number of variants: *ma*, *ma:*, *me*. These variants occur under various segmental and prosodic conditions. This morpheme is almost always written as a separate word, though there are a few cases in which it is joined to a following word. In Isa 3:15, *mlkm* is to be read *ma-lla:xém* lit. ‘what to you (m. pl.)’, (i.e., ‘how dare you’); in Ex 4:2, *mzh* is to be read *ma-zzé* ‘what is this?’ It is interesting that in both these cases the morpheme *ma-* is attached to a function word (a pronoun in the first case and a demonstrative in the second). Function words are more prone to be fused with other morphemes into single words.⁴

Thus, orthographic words are potential prosodic words. To qualify, a form must meet the minimum criterion of having at least a full syllable CV, where V is not schwa. Full vowels created by phonological processes do not count. On top of that, a morpheme must exhibit a certain syntactic-semantic independence – hence, *ma* is a potential prosodic word, *ha-* is not.

3. The Pointed Text: The Prosodic Word

The Biblical text was gradually stabilized and fixed in the centuries leading up to the first century of the common era (Cross and Talmon 1975; Sáenz-Badillos 1993). At a certain point, no further changes were permitted to be made to this text. Therefore, to this day Torah scrolls that are used for public readings consist only of a consonantal text, with no indications of verse divisions, stress, or other prosodic markers.

The prohibition against adding markings did not apply to texts intended for private or nonliturgical purposes, and symbols for vowels, consonant

diacritics, and an elaborate system of ‘accents’ to mark phrasing began to be introduced in the 6th and 7th centuries C.E., presumably to preserve the pronunciation of the traditional reading of the text (Goshen-Gottstein 1963). This activity was carried on for a number of generations by scholars known as Masoretes. A number of distinct but related schools arose; the best known was associated with a group working around the city of Tiberias, and so is known as the Tiberian system (Dotan 1971; Yeivin 1980).

The Tiberian system of accents represents a highly elaborated prosodic representation that, among other aspects of pronunciation, organizes the text into hierarchical groupings of verses, phonological phrases, and prosodic words (Dresher 1994). The Masoretes were not free to tamper with the consonantal text itself. While they could add diacritic marks over, under, beside, or even inside letters, they could not change or transpose letters, or add or remove spaces between words. To indicate that two or more orthographic words are to be considered as a single prosodic word, the Masoretes connected the words in question by a hyphen, called *maqgef*.

Whether an orthographic word in the consonantal text is cliticized to a following word or heads its own prosodic word has phonological consequences. A prosodic word has a single main word stress on the final or penultimate syllable. One rule that applies only to syllables bearing main word stress is Tone Lengthening:

(6) Tone Lengthening (Prince 1975)

Lengthen a vowel bearing main stress in its prosodic word.

Conditions:

- a. The rule does not apply to the low vowel /a/ when followed by two consonants;
- b. The rule does not apply to verbs.

This rule is exemplified by the accusative particle $\text{ʔe}\theta$. In the majority of cases, this particle is attached by *maqṣef* to the following word, indicating that it is cliticized to it and does not have its own word stress. In these cases the particle is pointed with the vowel *e*, as in (7a). When it is an independent prosodic word (7b), it is pointed with the vowel *e:*.

(7) The accusative particle

a. As clitic $\text{ʔe}\theta$ -

$\text{ʔe}\theta$ -ha:ʕó:r

acc-the.light (Gen 1:4)

b. As independent word $\text{ʔe}:\theta$

$\text{ʔe}:\theta$ hašša:máyim

acc the.heavens (Gen 1:1)

4. Cliticization in the Tiberian Text

Whether an orthographic word is cliticized or not depends on a complex set of prosodic, phonological, and syntactic conditions, some of which are reviewed in the following sections. It turns out that cliticization is tightly tied in with the entire Tiberian prosodic system, and cannot be understood without taking into account the principles of phrasing.

4.1. Rudiments of the Tiberian system of accents

For purposes of the current discussion, it is necessary to know that the Tiberian diacritics known as ‘accents’ fall into two groups. A conjunctive accent on a word indicates that the word is in the same phonological phrase as the word it follows; a disjunctive accent indicates that its word is phrase final. Disjunctive accents, in

turn, are arranged into four hierarchical classes, conventionally designated D0, D1, D2, and D3, where D0 represents the strongest disjunction (coming only at the end of a verse and at the end of the half-verse), and D3 represents the weakest (Cohen 1969). A phrase ending in an accent of level D_i is divided by an accent of level D_{i+1} , until the D3 level. A D3 phrase is divided by another D3. Therefore, unlike much contemporary work that assumes strict layering of phonological phrases, the Tiberian prosodic representation divides each verse into nested phonological phrases.⁵

The principles governing the division of a verse into phrases, and hence the distribution of the accents, are extremely complex, and though some of the leading principles and rules for particular circumstances are now known, much remains to be discovered (see Aronoff 1985; Breuer 1982; Drescher 1994; Janis 1987; Price 1990; Wickes 1887; among others). Cliticization is integral to the entire system, because phrasing is sensitive to the number of words and to the prosodic weight of words, and cliticization affects both: cliticization can change two short words into one long word, for example.⁶ Therefore, cliticization is woven into the phrasing algorithm; it cannot be regarded as a preliminary step that takes place prior to the division into nested phrases, or conversely, as a late fix up that follows the division of words into phrases.

The principles governing cliticization are therefore particularly complex, because, being situated at the interface between word and phrase, they involve general principles of phrasing as well as particular idiosyncrasies of lexical items. The most detailed discussion of cliticization in the Biblical text that I know of is that of Breuer (1982: Chap. 7). Breuer proposes a series of descriptive

generalizations that set out conditions under which cliticization is facilitated or blocked. These generalizations take the form of conditions akin to the ‘preference laws’ of Vennemann (1988) or the constraints of much current phonological theory, notably Optimality Theory (OT: Prince and Smolensky 1993). OT proposes a theory of how conflicts among constraints can be accommodated.

Generally speaking, the principal categories of cliticization are the following:

- (i) Small words: Some small words have an inherent tendency to be cliticized.⁷
- (ii) Simplification of phrasing: Cliticization simplifies the phrasing, either by reducing the number of conjunctive accents in a phonological phrase, or by reducing the number of phonological phrases.
- (iii) Clash avoidance: To avert a stress clash, by relieving the cliticized word of its clashing main stress.

4.2. Small words

Cliticization occurs most readily to small monosyllabic words that have a short vowel in a closed syllable. Breuer divides these words into two classes: those that are generally cliticized to any word, short or long, and those that are regularly cliticized only to short words. Breuer (1982: 167) gives the following list of words of the first class:

(8) Small function words that can be cliticized to any word

ʔeθ ‘accusative particle’, ʔal ‘on’, ʔel ‘to’, min ‘from’, ʔað ‘until’,
ʔim ‘with’, ʔim ‘if’, ʔal ‘not’, bal ‘not’, pen ‘lest’, ʔaf ‘also’, ma
‘what’, kol ‘all’, ben ‘son’, baθ ‘daughter’, ʔeθ ‘time’

Most of these words are straightforwardly function words – the accusative particle, prepositions, negative particles, various subordinating complementizers, and quantifiers. The nouns *ben* ‘son’ and *baθ* ‘daughter’ might appear to be content words; however, they are also used in contexts where their lexical meanings are attenuated or lost, and take on a more functional cast. The word *ben*, for example, can designate a quality (*ben-ḥáyil* lit. ‘son of valour’ = ‘valiant’), or mean ‘deserving of’ (*ben m:óθ* lit. ‘son of death’ = ‘he shall surely die’), or be part of an expression indicating age (*bəne:-ša:ná:* lit. ‘sons of a year’ = ‘of the first year’, i.e., ‘less than a year old’), and so on. Similar considerations apply to *baθ*, (e.g., *ʔéz baθ-šəna:θá:h* ‘a she-goat in its first year’, *baθ-bəliyyá:ʔal* lit. ‘daughter of baseness’, ‘a worthless woman’) and less obviously also to *ʔeθ* ‘time’, perhaps because of its association with time (*ləʔéθ ziqna:θó:* ‘in his old age’).

The interrogative *ma* ‘what’, though it fits semantically, appears to be out of place because it has an open syllable. However, Breuer points out that it functions as a closed syllable (*maC*) because it causes gemination of a following consonant. However, it is not cliticized when followed by a guttural, which does not geminate. Hence, we have the following pair:

- (9) Cliticization of *ma* when a geminating consonant follows

(ma-ppišŕí:)D2 (má ḥaṭṭa:θí:)D1

what-my.trespass what my.sin

‘What is my trespass? What is my sin?’ (Gen 31:36)

The first instance of the word *ma* in this verse is cliticized to a word with a geminating consonant. In the second phrase, the word following *ma* is long and has an initial consonant that is a nongeminating guttural; therefore, *ma* acts like a word ending in an open syllable, and does not cliticize in this phrase.

In addition to the small words in (8), Breuer identifies another set of small words that are more restricted in their tendency to cliticize. In general, these words tend to cliticize only to short words. Breuer divides these words into two lists: the words in the first list (10a) cliticize more readily than those in the second (10b).

- (10) Small (mostly) content words that can be cliticized to short words

- a. gam ‘also’, ḥax ‘but’, raq ‘only’, yaḏ ‘hand’, kaf ‘palm’, ŕam ‘people’, dam ‘blood’, davar ‘word (const.)’, har ‘mountain’, sar ‘officer’, gan ‘garden’, rav ‘great’, ḥay ‘holiday’, rax ‘soft’, nəḥum ‘speech’
- b. ḥaf ‘anger’, mas ‘tax’, gal ‘heap’, qaš ‘straw’, paḏ ‘morsel’, gaḏ ‘winepress’, bar ‘son (Aram.)’, ḥaḏ ‘one (Aram.)’, šen ‘tooth (const.)’, ḥoq ‘statute (const.)’, mor ‘myrrh (const.)’, tom ‘integrity (const.)’, tam ‘complete (const.)’, šal ‘doff’, raḏ ‘subdue’, ḥay ‘live’, ḥat ‘you (f. sg.)’, ze ‘this’, bah ‘in that (Aram.)’, bəḥaḏ ‘for’, nəqam ‘revenge (const.)’, šəḥar ‘young

of an animal (const.)', pəšar 'interpretation (const.) (Aram.)',
 ləven 'white (const.)', məloḥ 'reign'

These lists consist mainly of content words, though the first three words in (10a) are function words with syntactic/semantic values comparable to those in (8). Thus, it is not clear why these words are grouped with (10) rather than (8). Some of the content words in (10), like those in (8), also have wider uses that could arguably put them into the function word class. An example is *yaḏ* 'hand', which combined with various prepositions can mean 'by the side of', 'next to', 'at the disposal of', and so on. Note that *?af* 'anger', in (10b) is homonymous with *?af* 'also', in (8).

Besides the content-word function-word distinction, these words are also distinguishable phonologically. The final consonant in the noun *?af* derives from an underlying geminate which surfaces in suffixed forms, such as possessive *?appó:* 'his anger'. Many of the other words in (10) likewise have underlying final geminates; thus, these words are not just semantically 'heavier' than those in (8), but phonologically heavier, also, though the phonological distinction is neutralized at the surface in unsuffixed forms.

The nouns in (10a) differ from those in (10b) mainly in that the former are more common. The words in (10b) consist of Hebrew nouns together with a mixed bag of other parts of speech, including the odd verb, preposition, demonstrative, and pronoun, and even some words in Aramaic. Thus, tendency to cliticize depends on a variety of factors, including phonological weight, morphological/syntactic class, semantic function, and commonness.⁸

Many words in (10) are construct forms. The construct raises special problems for the definition of word, but we cannot pursue this topic here.

Breuer (1982: 168) points to pairs such as in (11) as showing how cliticization of words in group (10) is sensitive to the length of the following word; cliticization of *gam* ‘also’ applies before a short word (11a), but not before a long word (11b).⁹

(11) Cliticization of *gam* depends on the length of the following word

a. As clitic

(wəɣam-ʔo:θi:)D2

and.also-acc.me (2 Sam 2:7)

b. As independent word

(wəɣám ʔa:no:xí:)D2

and.also I (2 Sam 2:6)

Given what I have reported to here, one might expect that the words listed in (8) and (10) should be freely cliticizable – everywhere, in the case of the words in (8), and before short words, for (10). This, however, is not the case. As mentioned above, cliticization interacts with other aspects of phrasing. Other constraints on the phrasing algorithm can conflict with cliticization in certain situations, and in these configurations, cliticization is systematically blocked.

For example, there is a very strong constraint that the half-verse, which ends with a D0 accent, should consist of at least two phrases. In some verses, the main division is such that one of the half-verses contains only two words, one of which is a small cliticizable word. In such a case, the small word almost always remains an independent word in its own phrase, marked with a disjunctive accent.

(12) Half-verse contains only two words

(wəʔé:θ)D1 (bəθu:ʔé:l)D0

and.ACC Bethuel (Gen 22:22)

Another constraint that applies to D0 phrases is that a long word does not easily coexist with another word. This phenomenon can be understood as due to a slowing down of the reading in prominent positions, so that a long word in such a position counts as if it were two words, hence already enough to fill a whole phrase (Dresher and van der Hulst 1998). Thus, a small word is generally not cliticized to a long word in a D0 phrase, but again is placed in its own phrase with a disjunctive accent. This phenomenon is illustrated by the following verses:

(13) Cliticization of a small word to a long word

a. In a D1 phrase

(wəʔeθ-haggirga:ší:)D1

and.ACC-the.Girgashites

(Gen 15:21)

b. In a D0 phrase

(wəʔé:θ)D1 (haggirga:ší:)D0

and.ACC the.Girgashites

(Gen 10:16)

4.3. Simplification of phrasing

Cliticization can also occur to reduce the number of disjunctive accents, in order to create a smoother phrasing. In particular, expected phrasings of the form (14a) and (14b) below may be simplified as shown.

(14) Simplification

a. (w)D_{n+1} (w w)D_n ---> (w w-w)D_n

b. (w w)D_{n+1} (w)D_n ---> (w-w w)D_n

The likelihood of a word being cliticized in these contexts increases with increased shortness of the word. Of course, the small words in (8) cliticize to any following word unless blocked by the phrasing principles discussed in the previous section, so they would be cliticized in (14) without any further

stipulation. The small words in (10) cliticize as a matter of course only to short words; recall (11) above. The contrast in (11) arises where the conditions of (14) do not obtain; however, such words do cliticize to a long word when Simplification is possible:

(15) Cliticization of *gam* in Simplification context

a. (wəʔibba:né ɣam-ʔa:noxí:)D1

and.I.shall.be.built.up also-I (Gen 30:3)

b. *(wəʔibba:né)D2 (gám ʔa:noxí:)D1

Apart from the small words discussed above, cliticization applies most commonly, according to Breuer, to the subordinating complementizers *ki:* ‘that, for, when’, *ʔăšer* ‘that’, and the negative morpheme *lo:*. Cliticization of this type is illustrated in the following example:

(16) Example of Simplification via cliticization of *lo:*-

(wəʔyé:r lo:-θo:né)D1 (wəló: θilħa:şénnu:)D0

and.stranger not-vex and.not oppress.him

‘Thou shalt neither vex a stranger, nor oppress him’ (Ex 22:20)

The first instance of the word *lo:* in this verse is cliticized, thereby reducing the number of words from three to two and allowing the preceding word to be phrased in the same phrase, rather than forcing it into its own phrase. This is thus an example of (14a). The second instance of this word is not cliticized, because doing so would serve no simplifying purpose, nor is there a stress clash in this phrase. This example shows that, at least in the case of this lexeme, it is preferable to have two words in a phonological phrase than for it to be cliticized leaving just one word in the phrase, in the absence of other factors favouring cliticization.

Another form of Simplification is cliticization to reduce the number of words in a phrase by replacing a conjunctive accent. Such cases can be represented schematically as in (17).

(17) Reduction

a. (w w w)Dn ---> (w w-w)Dn

b. (w w w)Dn ---> (w-w w)Dn

The situations in (17) are the minimal ones in which Reduction can occur. That is, a word which is not inherently cliticizable by the criteria discussed above will not cliticize to reduce a phrase from two words to one word. Cliticization of this type can occur, however, in larger phrases, containing more than three words.

It follows that Reduction is relevant only in phrases ending in disjunctive accents that support more than one preceding conjunctive accent, that is, that allow more than two words in a phrase. As a rule, the less prominent the phrase, the more words can be fit into it. Therefore, Reduction situations arise often in D3 and D2 phrases, but only rarely in D1 or D0 phrases.

D0 phrases, for example, normally permit a maximum of two words. This generalization is never violated before a verse-final D0 accent. However, before the D0 that ends the first half-verse (an accent called *atnaḥ*), there is a particular situation in which more than two words can occur. Breuer (1982: 156) observes that this configuration arises when the word *ki:* is followed by a word with initial stress, which is in turn followed by the D0 word (18b). When *ki:* is followed by a word with noninitial stress in a comparable sequence, it is cliticized, resulting in an ordinary two-word D0 phrase (18a).

(18) Cliticization of *ki:* in a D0 phrase depends on following word

a. As clitic

(ki:-θe:lé:x ʔimmá:nu:)D0

rather-you.will.go with.us (Nu 10:32)

b. As independent word

(kí: vá: vilʔá:m)D0

that came Balaam (Nu 22:36)

The pattern exemplified in (18) is the opposite of what we might have expected; since words are cliticized as a way of averting a stress clash (see the following section), the failure of *ki:* to cliticize in just such a position is unexpected. Thus, (18b) is anomalous in two ways: the existence of a three-word phrase ending in D0, and the failure of cliticization. These two anomalies can be connected by supposing that *ki:* in (18b) is treated by the accent system as in fact being cliticized at some abstract level; for reasons that remain unclear, the cliticization is suspended in this particular configuration.

4.4. Clash avoidance

Cliticization can occur to prevent a stress clash between words in the same phonological phrase. In Tiberian Hebrew, a stress clash occurs between two words in the same phonological phrase when the first word has final stress and the second word has initial stress. If the first word ends in a superheavy syllable (a phonologically long vowel in a closed syllable), no clash is considered to occur.

The cliticized word in (19a) has final stress when independent (19b). The effect of cliticization is to deprive the word of its main word stress, thus averting a

stress clash. Notice that a secondary stress appears on the initial syllable of the cliticized word in (19a). This secondary stress arises through the ordinary rule of secondary stress assignment, counting back two full syllables from the main stress of the entire prosodic word (Dresher 1981b):

(19) Cliticization to avert a stress clash

- | | |
|-------------------------|-------------------------|
| a. (vàyhi:-f'érev...)D1 | b. (vayhí: va:f'érev)D1 |
| and.was-evening | and.was in.the.evening |
| ‘And there was evening’ | ‘When evening came’ |
| (Gen 1:8) | (Gen 29:23) |

Another means of averting a stress clash is by stress retraction (McCarthy 1979; Rappaport 1984; Revell 1987), examples of which are shown in (20); the first word in each phrase is normally stressed on the final syllable.

(20) Stress retraction to avert a stress clash

- | | |
|---------------------------|-------------------------------|
| a. (qá:ra: lá:yla:)D0 | b. (tó:xal léhem)D1 |
| he.called night (Gen 1:5) | you.will.eat bread (Gen 3:19) |

The vowel onto which stress is retracted must normally be long. For words that do not meet the conditions for stress retraction, cliticization is the only option for avoiding a stress clash, as in (19a). Where retraction can occur, it appears to be the preferred option; where retraction is not permitted, cliticization can occur. Sometimes the stress clash is left unresolved, for reasons explained in detail by Revell (1987). An example illustrating these two options side by side in the same phonological phrase is the following:

(21) Stress retraction and cliticization to avert stress clashes

(mó:zne: šéðeq ʔavne:-šéðeq)D2

balances honest weights-honest

‘an honest balance, honest weights’ (Lev 19:36)

In the word *mó:zne:*, stress can retract onto the phonologically long vowel *o:*; but the initial vowel of *ʔavne:* is short, so cliticization is the only available option (short of leaving the clash unresolved).

4.5. Summary

The preceding remarks on cliticization present only a partial picture of this complex phenomenon. They should suffice to show, however, that the principles governing the distribution of prosodic words in the text are bound up with constraints on phrasing that operate at higher levels of the prosodic hierarchy. Earlier proposals for mapping the prosodic structures indicated by the accents form the syntax assumed a derivational approach, whereby prosodic structure is built up in a series of steps (Dresher 1994; Janis 1987). The above survey suggests that evaluation of candidate forms by means of ranked constraints, as proposed by OT, offers a promising alternative. I will not, however, attempt such an analysis here (though the reader is invited to begin to construct one from the materials presented above).

5. The Word Level in Phonology

Up to now we have considered the notion of the word as represented orthographically in the Tiberian text, and have found two types of words associated with the consonantal and pointed text, respectively. Both notions relate

to the prosodic word of linguistic theory. The orthographic word of the consonantal text corresponds to potential prosodic words, that is, words that can stand as independent prosodic words in some context. The pointed text indicates which of these potential prosodic words are actually realized as such and which are cliticized.

There is another notion of ‘word’ that is relevant to phonological theory, in the sense of a level at which certain phonological and morphological processes apply. In the theory of Lexical Phonology and Morphology (LPM: Kiparsky 1982, 1985; Mohanan 1982, 1986; Pesetsky 1979), phonology and morphology apply in stages to a series of levels, such as the stem, the word, and postlexical levels. Though a prosodic word is necessarily a domain for word-level processes in LPM, the word level of LPM is not exhaustively characterized by the prosodic words indicated in the Tiberian pointed text. That is, word-level processes apply also to certain constituents of prosodic words.

In Dresher 1983, I argued that Biblical Hebrew displays some level ordering, but the levels are not exactly what we might expect from the results of other studies. In brief, I argued that there is very little evidence for stem-level phonology, apart from some minor rules that apply to particular morphemes. I connected this fact to the nonconcatenative nature of much Semitic morphology, which does not create a suitable environment for the operation of typical phonological processes. Thus, there is no evidence that suffixes, for example, need to be distinguished as being stem-level or word-level suffixes. Unlike many dialects of Arabic (Brame 1974; Broselow 1976; Kenstowicz 1981; Kiparsky 2002b), for example, object suffixes, though attached outside of subject-

agreement suffixes as in other Semitic languages, do not appear to be attached at a different level than other suffixes. For purposes of stress, syllabification, and segmental phonology, a word like *yišmorxá*: ‘he will guard you’, from underlying /ya+šmor+e+ka/, can be subjected to affixation and word-level phonology in one pass, without any internal cycles. It follows from LPM that we should not expect the morphology to be sensitive to derived phonological properties, and this prediction is borne out; unlike English, for example, Biblical Hebrew has no word-formation rules that make reference to stress.

I also argued that the word level itself must be conceived of differently than in most studies of LPM. We can think of a word as existing paradigmatically in the lexicon, or else as being syntagmatically placed within a phrase. In most studies of LPM, the word is thought of, perhaps correctly, as being in the lexicon. Hence, languages like English have category-changing word-level affixes that are best thought of as being assigned in the lexicon prior to insertion in the syntax (e.g., *sing#er*, *sad#ness*, *national#ize*). Consistent with this is the fact that word-level phonology in English is not sensitive to the position a word has in its phrase.

In Biblical Hebrew, however, both these phenomena point in a different direction: here, word-level phonology applies not to words in the lexicon, but to words already placed in a phrase. Evidence for this is that word-level phonology in Hebrew is sensitive to the position of a word in a phrase. This evidence comes from the so-called pausal forms, a ubiquitous feature of Tiberian Hebrew prosody. In Tiberian Hebrew, many words have one form when they are phrase-internal – the contextual form – and another form, called the pausal form, when they are final in a major phrase (which I take to be the intonational phrase of the

contemporary prosodic hierarchy – see DeCaen 2005; Drescher 1994; Goerwitz 1993; Revell 1980, 1981). In most cases, both the contextual and pausal forms can be derived from a common underlying source by the same regular rules of the phonology. The source of the difference can be located in the way rules of stress and reduction apply in pause and in context. Starting from /ya+šmor+e+ka/, for example, the penultimate vowel is reduced when the word is in context, causing stress to appear on the final vowel, and preserving the stem vowel, the result being *yišmorxá:* as cited above. When the same word is in pause, however, the penultimate vowel is retained and stressed, resulting in the reduction of the stem vowel, yielding *yišmərə́xa:* (see Malone 1993; Prince 1975 for details). If word-level phonology waits until the whole word has been put together and inserted into its phrase, it follows from LPM that there should be no word-level category-changing affixes, and this appears to be correct.

Although there exist no word-level affixes that apply in the lexicon, there does exist a class of word-level prefixes that create a word-level cycle in the phrase:

- (22) Word-level prefixes
- a. Prepositions: b(ə)- ‘in’, l(ə)- ‘to’, k(ə)- ‘like’, mi- ‘from’
 - b. Conjunctive: w(ə)- ‘and’
 - c. Definite article: ha- ‘the’

The prefixes in (22) are just those discussed in §2 as being noteworthy in that they are not written as independent words in the consonantal text, though their syntactic-semantic status might qualify them as being grammatical words. In the case of the prepositions (22a), we observed that they each have variants or

synonyms that are independent orthographic words. From the point of view of the syntax, too, these prepositions are best viewed as being introduced into the syntax as independent morphemes with their own syntactic positions. They are then obligatorily cliticized, but not before certain word-level phonological processes have applied.

One of these is a rule changing /a/ to *i* in a word-initial closed syllable:

(23) *A-to-I* (Prince 1975)¹⁰

a → i/#C_____CC

This rule applies to (24a), where there is no prefix. The underlying /a/ surfaces in prefixed forms, like (24b). However, the rule applies despite the presence of one of the prefixes in (22), as shown in (24c).

(24) Examples of *A-to-I*

a. No prefix: *A-to-I* applies

/gaddel/ → giddé:l ‘he brought up’

b. Lexical prefix: *A-to-I* does not apply

/ya+gaddel/ → yəgaddé:l ‘he will bring up’

c. Word-level prefix: *A-to-I* applies

/w(ə)#gaddel/ → wəgiddé:l ‘and he brought up’

Another phenomenon that attests to the distinctive status of the word-level prefixes is spirantization. Spirantization normally applies to a (nonemphatic and nongeminate) stop that immediately follows a vowel (see Idsardi 1998 for detailed discussion). Hence, in (25a) the /k/ and /b/ of the root /ktb/ are spirantized, but the /t/ is not. Following a word-level prefix, however, the /t/ is also spirantized (25b).¹¹

(25) Contrast in Spirantization

- a. Lexical prefix: Medial C of root not spirantized

/la+ktob/ → lixtó:v ‘to write’

- b. Word-level prefix: Medial C of root spirantized

/b(ə)#ktob/ → bixθó:v ‘when writing’

This spirantization attests to the presence of a vowel between the *k* and the *t* in (25b). Such a vowel would arise on the inner cycle in /ktob/ to break up the illicit initial consonant cluster, spirantizing the /t/ (26b). On the outer cycle, the prefix vowel (whether underlying or inserted to break up the CC cluster) spirantizes the /k/. The addition of the prefix also puts the derived vowel of the inner cycle in the context VC_CV; this is a context in which short vowels are deleted, hence the output *bixθó:v*. In (26a), the prefix plus stem form a single word domain, and so there is no reason to insert a vowel between the first two root consonants (Borer 1979; Idsardi 1998; Rappaport 1984).

(26) Word domains

	a. Lexical prefix	b. Word-level prefix
Inner cycle	la-ktob	ktob
Input	(laktob) _w	(ktob) _w
Output	(lixtó:v) _w	(kVθó:v) _w
Outer cycle		b-(kVθó:v) _w
Input	---	(b(kVθó:v) _w) _w
Output	---	bixθó:v

In sum, Biblical Hebrew supports the general picture of phonology-morphology interaction posited by LPM. It also shows that word-level phonology

may have access to the position of a word in a phrase, and that there exist postlexical levels that have many of the properties of lexical levels (see Dresher 1983 for further discussion).

Of course, since phrasing itself depends on some derived phonological properties (notably, the position of stress), the nature of the interaction between level-ordered phonology and the phrasing algorithm is not entirely clear. Putting together the results of this section and the previous one, it appears that an adequate analysis of Biblical Hebrew phonology and prosody may require a derivational component as well as parallel constraint evaluation, perhaps along the lines sketched by Kiparsky (2002b).¹²

6. Conclusion

For hundreds of years, Biblical Hebrew has been at the centre of important developments in linguistics. In this brief survey of some aspects of the word in Tiberian Hebrew I hope to have shown that this position is entirely merited, and that the Masoretic text continues to raise interesting and complex problems that are relevant to current issues in linguistic theory.

Notes

* It is an honour and a pleasure to dedicate this article to Paul Kiparsky, who has contributed so much to our understanding of the word and its place in phonology and morphology. For various kinds of valuable help and illuminating discussions of Biblical Hebrew, I would like to thank Jean Balcaen, Vincent DeCaen, and Bill Idsardi. This research was supported in part by Social Sciences and Humanities Research Council of Canada research grant 410-96-0842.

1. Joüon (1947: §8), against the opinion of Kautzsch (Bergsträsser 1962; Gesenius 1910), considers that a stem-initial schwa was actually pronounced, though in a weakened form that does not amount to a normal reduced syllable: *bið^əvar*. His arguments for this assumption are first, that the schwa corresponds to a vowel that was historically present, and second, that the rule of spirantization applies to the following consonant, indicating the presence of a vowel (**biðbar*). However, neither of these arguments is compelling: the historical existence of a vowel does not necessarily bear on its synchronic status; and though spirantization does point to the synchronic presence of a vowel, it does not necessarily indicate that this vowel is present at the surface - see Idsardi (1998) and §5.

2. A remark is needed concerning the transcription of vowels used here and the issue of vowel quantity. The Tiberian transcription distinguishes seven vowel signs, and the current consensus is that these vowels are distinguished by quality, not quantity, with values approximating to [i, e, ε, a, ə, o, u] (Bergsträsser 1962; Joüon 1947; Khan 1987). Despite the apparent seven-vowel system of the

Tiberian transcription, there is a long-standing tradition (Chomsky 1952) of considering the underlying vowel system of Biblical Hebrew to comprise ten vowels, symmetrically divided into five long and five short: /i:, i, e:, e, a:, a, o:, o, u:, u/. Even while indicating vowel quantity in transcriptions, writers such as Joüon 1947 and Lamdin 1971 are noncommittal as to the phonetic reality of this scheme. However, there is no doubt that the quantitative interpretation makes much better sense of the phonological alternations of Biblical Hebrew than does the purely quantitative interpretation. Thus, I will refer to lengthening and long and short vowels, understanding these terms to refer to a genuine phonological reality in the grammar of Tiberian Hebrew, though not necessarily at the surface phonetic level.

3. I omit other variants of the definite article that arise in various environments having to do with the position of stress and other peculiarities of the gutturals.
4. For example, prepositions cannot occur with independent pronoun forms, but only with suffixal forms of the pronoun: *ləxá:* or *?e:léxa:* ‘to you’, never **?él ?attá:*, where *?attá:* is the independent form of the 2nd person masculine singular pronoun.
5. See Drescher 1994 for further discussion of the rationale behind this nesting and its connection with contemporary approaches to prosodic structure.
6. A long word has at least two full syllables before the main stress; a short word does not meet this condition. See Drescher 1981a for discussion of the theoretical basis underlying these definitions.
7. A small word is a word with only one syllable (not counting schwa).

8. Breuer (1982: 171) writes that he includes *nəʔu:m* ‘speech’ in this list even though it has a long vowel in a closed syllable because it cliticizes frequently, particularly in certain fixed phrases. *ze* ‘this’ appears to be out of place here because it ends in an open syllable; however, like *ma* ‘what’ it causes gemination of a following consonant when it cliticizes, thereby closing its syllable.
9. The initial /g/ is spirantized after a vowel in both examples in (11).
10. See Prince (1975: 157) for a refinement of this rule.
11. Note that the infinitive prefix *la+* is distinct from the preposition *l(ə)#*.
12. See also Dresher (in press) for further discussion of this issue.

References

- Aronoff, Mark. 1985. Orthography and Linguistic Theory: The Syntactic Basis of Masoretic Hebrew Punctuation. *Language* 61: 28–72.
- Bergsträsser, G. 1962. *Hebräische Grammatik, mit Benutzung der von E. Kautzsch bearbeiteten 28. Auflage von Wilhelm Gesenius' hebräischer Grammatik. Mit Beiträgen von M. Lidbarski*. Hildesheim: Georg Olms. [Reprint of the edition published in Leipzig, 1918.]
- Borer, Hagit. 1979. Schwas in Tiberian Hebrew: An Argument for Cyclic Rules of Syllable Structure Modification. *MIT Working Papers in Linguistics* 1: 149–171.
- Brame, Michael. 1974. The Cycle in Phonology: Stress in Palestinian, Maltese and Spanish. *Linguistic Inquiry* 5: 39–60.
- Breuer, Mordecai. 1982. *Tašame hamiqra bekhaf-ʔalef sefarim uvesifre ʔemet* [The Biblical Accents in the Twenty-One Books and in the Three Books]. Jerusalem: Mikhlala.
- Broselow, Ellen. 1976. The Phonology of Egyptian Arabic. Doctoral dissertation, University of Massachusetts, Amherst.
- Chomsky, William. 1952. *David Kimḥi's Hebrew Grammar (Mikhlol). Systematically Presented and Critically Annotated*. New York: Bloch.
- Cohen, Miles B. 1969. *The System of Accentuation in the Hebrew Bible*. Minneapolis: Milco Press.
- Cross, Frank Moore, and Shemaryahu Talmon (eds.) 1975. *Qumran and the History of the Biblical Text*. Cambridge, Mass.: Harvard University Press.

- DeCaen, Vincent. 2005. On the Distribution of Major and Minor Pause in Tiberian Hebrew in the Light of the Variants of the Second Person Independent Pronouns. *Journal of Semitic Studies* 50: 321–327.
- Dotan, Aron. 1971. *Masorah*. In *Encyclopædia Judaica* 16: 1401–1482.
- Dresher, Bezael Elan. 1981a. Accentuation and Metrical Structure in Tiberian Hebrew. *MIT Working Papers in Linguistics* 3: 180–208.
- Dresher, Bezael Elan. 1981b. Metrical Structure and Secondary Stress in Tiberian Hebrew. *Brown University Working Papers in Linguistics* 4: 24–37.
- Dresher, Bezael Elan. 1983. Postlexical Phonology in Tiberian Hebrew. *Proceedings of the Second West Coast Conference on Formal Linguistics (WCCFL)*. Stanford: CSLI Publications. 67–78.
- Dresher, Bezael Elan. 1994. The Prosodic Basis of the Tiberian Hebrew System of Accents. *Language* 70: 1–52.
- Dresher, Bezael Elan. in press. Stress Assignment in Tiberian Hebrew. In Charles Cairns and Eric Raimy (eds.), *Architecture and Representations in Phonological Theory*. Cambridge, MA: MIT Press.
- Dresher, Bezael Elan & Harry van der Hulst. 1998. Head-dependent Asymmetries in Phonology: Complexity and Visibility. *Phonology* 15: 317–352.
- Elliger, Karl, and Wilhelm Rudolph (eds.) 1977. *Biblia Hebraica Stuttgartensia*. Stuttgart: Deutsche Bibelstiftung.

- Gesenius, Wilhelm. 1910. *Hebrew Grammar, as Edited and Enlarged by the Late E. Kautzsch*. 2nd English edition revised in accordance with the 28th German edition (1909) by A. E. Cowley. Oxford: University Press.
- Goerwitz, Richard. 1993. *Tiberian Hebrew Pausal Forms*. PhD dissertation, University of Chicago.
- Goshen-Gottstein, M. H. 1963. The Rise of the Tiberian Bible Text. *Biblical and Other Studies*, ed. Alexander Altmann, 79–122. Cambridge, Mass.: Harvard University Press.
- Idsardi, William J. 1998. Tiberian Hebrew Spirantization and Phonological Derivations. *Linguistic Inquiry* 29: 37–73.
- Janis, Norman. 1987. *A Grammar of the Biblical Accents*. Doctoral dissertation, Harvard University.
- Joüon, Paul. 1947. *Grammaire de l'hébreu biblique*. Deuxième édition anastatique corrigée. Rome: Institut Biblique Pontifical.
- Kenstowicz, Michael. 1981. A Note on Cyclic Stress in Levantine Arabic. *MIT Working Papers in Linguistics* 3: 21–31.
- Khan, Geoffrey. 1987. Vowel Length and Syllable Structure in the Tiberian Tradition of Biblical Hebrew. *Journal of Semitic Studies* 32: 23–82.
- Kiparsky, Paul. 1982. Lexical Morphology and Phonology. *The Structure of Phonological Representations*, Part 1, ed. Harry van der Hulst and Norval Smith, 131–175. Dordrecht: Foris.
- Kiparsky, 1985. Some Consequences of Lexical Phonology. *Phonology Yearbook* 2: 85–138.

- Kiparsky, Paul. 2002b. Paradigmatic Effects and Opacity. Ms., Department of Linguistics, Stanford University.
- Lambdin, Thomas O. 1971. *Introduction to Biblical Hebrew*. New York: Charles Scribner's Sons.
- Malone, Joseph L. 1993. *Tiberian Hebrew Phonology*. Winona Lake, Ind.: Eisenbrauns.
- McCarthy, John J. 1979. Formal Problems in Semitic Phonology and Morphology. Doctoral dissertation, MIT.
- Mohanan, K. P. 1982. Lexical Phonology. Doctoral dissertation, MIT.
- Mohanan, K. P. 1986. *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- Pesetsky, David. 1979. Russian Morphology and Lexical Theory. Ms., Department of Linguistics, MIT.
- Price, James D. 1990. *The Syntax of Masoretic Accents in the Hebrew Bible* (Studies in Bible and Early Christianity, 27). Lewiston, N.Y: Edwin Mellen.
- Prince, Alan. 1975. The Phonology and Morphology of Tiberian Hebrew. Doctoral dissertation, MIT.
- Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical report, Center for Cognitive Sciences, Rutgers University, New Brunswick, N.J., and Computer Science Department, University of Colorado, Boulder.
- Rappaport, Malka. 1984. Issues in the Phonology of Tiberian Hebrew. Doctoral dissertation, MIT.

- Revell, E. J. 1980. Pausal Forms in Biblical Hebrew: Their Function, Origin and Significance. *Journal of Semitic Studies* 25: 165–179.
- Revell, E. J. 1981. Pausal Forms and the Structure of Biblical Poetry. *Vetus Testamentum* 31: 186–199.
- Revell, E. J. 1987. *Nesiga (Retraction of Word Stress) in Tiberian Hebrew*. (Textos y estudios ‘Cardenal Cisneros’.) Madrid: Instituto de Filología, C.S.I.C.
- Sáenz-Badillos, Angel. 1993. *A History of the Hebrew Language*. Translated by John Elwolde. Cambridge: Cambridge University Press.
- Vennemann, Theo. 1988. *Preference Laws for Syllable Structure and the Explanation of Sound Change. With Special Reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.
- Wickes, William. 1887. *Two Treatises on the Accentuation of the Old Testament*. Reprinted with a Prolegomenon by Aron Dotan. New York: Ktav, 1970.
- Yeivin, Israel. 1980. *Introduction to the Tiberian Masorah*. Translated and Edited by E. J. Revell (Masoretic Studies 5). Missoula: Scholars Press.