The Behavior of Strong and Weak Verbs in Modern and Tiberian Hebrew: An OT Account

Joel M. Hoffman
Hebrew Union College
JMH@Huc.Edu / JMH@Exc.Com

Abstract

An Optimality-Theoretic (OT) account of verbal templatic morphology and phonology in both Tiberian Hebrew (TH) and Modern Hebrew (MH) is presented. This account is unique for three reasons: it provides a consistent explanation of both strong and weak verbs, including biliteral-root verbs; it provides an explanation of how TH became MH, showing in particular how a reranking of PARSE produced a massive repatterning of verbs in MH; thirdly, one possible stimulus for the constraint reranking is discussed, thus showing not only how but why TH became MH. Due to space requirements, only the first of these three is discussed in detail.

The current model further differs from traditional accounts in that it assumes that syllabification requirements comprise formatives in the verbal paradigms.

1 Introduction

1.1 Background

Hebrew is similar to Arabic in that much of its morphology derives from autosegmental representation of formatives solely on a consonantal tier, so that base verbs (to which conjugational affixes may be added) are, apparently, formed by putting consonantal roots into C-V templates. By way of example, from the consonantal root q.t.l we find the Hebrew verb *qatal* 'he killed,' on a par with the Arabic *qatala,* with the same meaning. The Hebrew is ostensibly from the template CaCaC, and the Arabic from the template CaCaCa.

Sometimes, a template will have more consonant slots than the root has consonants, in which case, it seems, one consonant appears more than once. For example, from the Arabic template

CaCCaCa we get *qattala* 'he massacred.' Hebrew has a similar form *qittal.*[1]

To these base forms, affixes are added for conjugation. For example, the 2-m-pl ending in Arabic is *-tum* and in Hebrew *-tem.* So we find *qataltum* 'you killed' and *qattaltum* 'you massacred' in Arabic, and *qataltem* and *qittaltem* in Hebrew.

In Arabic, verbs with identical second and third radicals behave as though they were biliteral, inasmuch as they follow the same sorts of patterns for spreading consonants seen when triliteral roots are put into templates with four consonantal positions. For example, from the root r.d.d in Arabic we find *radda* 'he replied.'

Based on these observations, McCarthy (1981) formulated his well-known **Obligatory Contour Principal** (OCP), according to which information on a single tier must not be repeated. His OCP does not allow a root such as r.d.d, and so the root must in fact be r.d, which explains why it patterns as though one radical were spreading. His OCP, a few spreading rules, and his templates account nicely for (almost)[2] all of the Arabic data. In particular, given a base form, the usual pronominal endings apply in the usual ways. So, for example, *radadtum* 'you replied.'

1.2 The Problem

Roughly equivalent to the the Arabic *radda* we find the (somewhat obscure) Hebrew word *rad* 'he repelled/he subdued' and its variant form *radad.* However, in Tiberian Hebrew, when *-tem* is added, we find *radotem,* and not **radadtem* or **radtem.* To make matters worse, Modern Hebrew has replaced *radotem* with *radadtem.*

It is these data that are to be explained. Specifically, why do these weak verbs pattern the way they do? And why have their forms changed?

In fact, evidence abounds suggesting that in general Hebrew verbs are not formed from templates. Forms such as *sovev* 'he spun' from the putative CiCCeC template and s.b, *nolad* from niCCaC and y.l.d, and *havinoti* from hiCCaCti and b.y.n demonstrate amply. Further, OCP violations are attested, as in as the roots *m.m.š* and *'.r.b.b.*

1.3 "Hebrew"

"Hebrew" refers at once to Tiberian Hebrew (TH) and to Modern Hebrew (MH). The former consists mostly of data reconstructed from the work of Ben Asher and his family, concluded around the ninth century, while the latter is the spoken Hebrew of the modern State of Israel. While a speaker of the MH can read and understand TH, significant changes in verbal morphology separate the two. Any theory that assumes that the TH paradigm is the only possible one cannot account

[1] More often, for unclear reasons, *qittel.*

[2] There are roots in Arabic that violate the OCP: *babGaa'* ('parrot') for example.

for MH, and vice versa. The theory presented here shows how the two are actually quite similar in spite of massive surface differences in some verbs. Specifically, the re-ranking of PARSE in MH will be shown to be responsible for the changes.

Biliteral roots in particular have changed considerably between TH and MH.

The current work attempts to provide an explanation of the more common triliteral roots and the biliteral roots, in both TH and MH, showing that the seemingly unexpected behavior of the weak roots can be predicted from an improved model of Hebrew morphology and phonology.

I assume the OT framework of Prince & Smolensky (to appear) etc.

In the spirit of Bat-El (1989), I assume that the verbal paradigms (see immediately below) in TH and MH cannot simply be C-V templates. Sharvit (1994), who also discusses MH morpho-phonology in an OT framework, assumes that each verbal form has associated with it a vowel tier, and that the final verb forms result from the combination of optional affixes (indicating agreement), that vowel tier, and the root consonants. I argue here, however, that many of the vowels that she assumes must be specified as part of the vowel tier follow from other considerations. Further, Sharvit (1994) does not discuss weak forms, in which the vowels change. The account here predicts those changes.

So, rather, I assume that the verbal paradigms are the result of consonantal roots and syllable requirements, in addition, perhaps, to vowels. That is, each paradigm (also called a *binyan*), is specified not only by its vowels but by a vowel and a syllable shape. More specifically, I assume that each paradigm is in fact a set of constraints on which vowels, prefixes, and syllables must appear in a base; suffixes and prefixes are then added to the base.

In this sense, too, the present work is not just a translation of old technology into new, but a significant deviation in terms of the underlying nature of Hebrew paradigms.

2 Data

Unfortunately, the data are complex, particularly for those not familiar with Semitic languages. Although they are reviewed elsewhere, very few descriptions capture both TH and MH, and especially few of them consider weak verbs. For these reasons, I present here an overview of Tiberian and Modern Hebrew. The reader familiar with these data may wish to skip ahead, and even the reader unfamiliar with them may wish to do the same after a quick perusal of the information below, as much of this information is repeated as necessary throughout this article.

For our purposes here, we will only consider the past tense of three paradigms, as they serve to illustrate most major features of the Hebrew verbal system.

In addition to the following descriptions, Table 1 will prove helpful.

Table 1: Summary of Hebrew Data

Paradigm		*Pa'al (Kal)*			*Niphal*			*Piel*		
Root		*3-m-s*	*3-pl*	*2-m-pl*	*3-m-s*	*3-pl*	*2-m-pl*	*3-m-s*	*3-pl*	*2-m-pl*
q.t.l	TH	qatal	qat'lu	q'taltem	niqtal	niqt'lu	niq'taltem	qitel	qit'lu	qitaltem
	MH	qatal	qat'lu	qataltem	niqtal	niqt'lu	niq'taltem	qitel	qit'lu	qitaltem
š.b.r	TH	šavar	šav'ru	š'vartem	nišbar	nišb'ru	nišbartem	šiber	šib'ru	šibartem
š.b.r	MH	šavar	šav'ru	šavartem	nišbar	nišb'ru	nišbartem	šiber	šib'ru	šibartem
(š.v.r?)					nišvar	nišv'ru	nišvartem	šiber	šib'ru	šibartem
k.r.s.m								kirsem	kirs'mu	kirsamtem
q.l	TH	qal	qalu	qalotem	naqal	naqalu	n'qalotem	qolel	qol'lu	qolaltem
	MH	qal	qalu	qaltem	naqal	naqalu	naqaltem	qolel	qol'lu	qolaltem
s.b	TH	sav	sabu	sabotem	nasav	nasabu	n'sabotem	sovev	sov'vu	sovavtem
(s.b.b?)		savav	sav'vu							
s.v	MH	sav	savu	savtem	nasav	nasavu	nasavtem	sovev	sov'vu	sovavtem
s.v.v	MH	savav	sav'vu	savavtem						
h.l.l								hilel	hil'lu	hilaltem
k.l.l	MH	kalal	kal'lu	kalaltem	nixlal	nixl'lu	nixlaltem	kilel	kil'lu	kilaltem
f.k.s.s	MH							fikses	fiks'su	fiksastem
f.k.s	MH							fikes	fik'su	fikastem
m.m.š	MH							mimeš	mim'šu	mimaštem
q.l.q.l								qilqel	qilq'lu	qilqaltem
'.r.b								'irev	'ir'vu	'iravtem
'.r.b.b								'irbev	'irbavtem	'irb'vu

2.1 Spirantization

The issue of spirantization in Hebrew is complex. In TH, descriptively, the sounds /bgdkpt/ spirantize and the end of a word, and after a vowel unless having once been geminate. For example, from *liktob* we have *lixto*β. For ease of reading and to facilitate comparison among forms, the symbols "θ," "ð," "γ" and "β" will be used sparingly, with *t,d,g,v,* respectively, used in their stead when the spirantization is immaterial.

Sometimes the vowel causing this spirantization disappears, for reasons having to do with word stress and vowel deletion. For example, from *katabu* we have *ka*θβ*u.*[3]

In TH the spirantized versions of these sounds do not appear morphemically.

In MH, on the other hand, the situation is chaotic. The sounds /gdt/ never change, leaving only /bkp/.[4] In base words (i.e., words without suffixes or prefixes) those tend to spirantize in the same places as they would in TH, but exceptions abound, often according to idiolect.[5] For example, both *safarti* and *saparti* 'I counted' are attested.[6] Affixes almost never cause a sound to spirantize in MH, even when they would in TH.[7] For example, the clitic prefix *l* 'to' when attached to *bayit* in TH results in *l'vayit,* but in MH we generally find *l'bayit.*

Additionally, MH has developed morphemic spirantized sounds /vxf/.

The situation is perhaps best seen by example. From the root χ.b.r 'join' in TH come the words *hit*χ*aber* 'became joined' and χ*aver* 'friend/member.' In TH, of course, the word χ*aver* could come only from the root χ.b.r. In MH, it has been reanalyzed as coming from the root χ.v.r. Accordingly, "became friends" is, in MH, *hit*χ*aver*, contrasting minimally with *hit*χ*aber,* which still exists with the meaning "became joined," and which is the prescriptive form for "became friends."

Other cases of morphemic /fvx/ abound in MH.

---

[3]Some researchers have suggested that the proper pronunciation of the form is *ka*θəβ*u.* But the only evidence for this is the spirantization of the consonant following the alleged shewa, and the erroneous assumption that the spirantization must result from surface vowels. Syllabification evidence (poetry etc.) indicates that the word is bisyllabic, a fact which even those who advocate the shewa admit.

[4]The reader will immediately notice that /bkp/ does not form a natural class.

[5]Even in TH, occasional exceptions are found. For example, both *birxa*θ and *birka*θ 'blessing of' are attested. It is not clear if this reflects underlying linguistic judgements from the time, or merely transcription errors which have became canonized in the relevant texts.

[6]The former is the prescribed form.

[7]One notable exception to this pattern is the prescriptive Hebrew used for radio broadcasts, in which the classical TH rules are (artificially) enforced. A discussion of the possible influences this dialect might have on the spoken language, particularly in a country in which radio broadcasts are an integral part of daily life, is beyond the scope of this article, but surely a source for interesting investigation.

2.2 Paradigms

For didactic reasons, spirantization will be partially ignored in the data that follow, thus making the nature of the system clearer.

Hebrew verbs come in (at least)[8] seven verbal paradigms, each of which roughly indicates a meaning variation.[9] For example, *katav* is "he wrote," *kitev* "he dictated," *niktav*[10] "[he/it] was written," *hitkatev* "he corresponded [by writing]." All of these share the root k.t.v, with various other patterns contributing to the meanings of the words.

Looking at three of the paradigms will suffice to see most of the interesting behavior. We will not look at the passive paradigms *huphal* and *pual* for reasons of space. The *hitpael* patterns with *piel*, and *hiphil* more or less with *niphal.*

The paradigms are traditionally described as consisting of C-V templates (McCarthy (1981)) or vowel patterns (Bat-El (1989)); here I assume they are syllabification requirements. At any rate, the final result is that each paradigm depicts a typical C-V pattern, with the pattern dictated by the paradigm and the choice of C's determined by the verbal root. In other words, the choice of consonants does not in general dictate the shape of the final verb.[11]

The simplest form in every case in the past-tense[12] third-person masculine. In the past-tense, suffixes are added to this form to create other persons and numbers. In the future tense, both prefixes and suffixes are used.

The affixes do not vary across paradigms, but are different in different tenses.

2.2.1 Kal

The simplest paradigm is called *pa'al*,[13] or *kal.* It generally appears CVCVC, as in *katáv* 'he wrote' from k.t.b or *šamár* 'he kept' from š.m.r. I assume here that the defining characteristic of *kal* is that the base's final syllable must be bimoraic. In the case of the third person past (which we have just seen) the base is the entirety of the final word.

Other declensions in the past tense are obtained by adding suffixes, which also tend to change the vowel patterns, according to rules too complex to describe here. These are the affixes that we will need for this our current purposes: *-ti* 'I,' *-tem* 'you pl.,' and *-u* 'they.' So: *katávti* 'I wrote,' *k'tavtém* (TH) or *katávtem* (MH), and *katvú.*[14]

[8]There are other marginal paradigms best ignored for now.

[9]The extent to which the paradigms dictate meanings is not clear, but for our purposes is irrelevant.

[10]On the surface, *nixtav.*

[11]Except when those consonants are weak, that is, either semi-vowels or sometimes /n/.

[12]Or perfective.

[13]Unlike in Arabic, where the paradigms are traditionally numbered, in Hebrew they are named according to how the root p.'.l would appear.

[14]This is sometimes *katávu* in TH.

2.2.2 Niphal

The passive of the *kal* is often the *niphal.* It generally appears as niCCaC, as in *nišmar* 'was kept' from š.m.r. It, too, is typified by a bimoraic final syllable in the base, and also by the prefixed /n/ on the surface.

The affixed forms we need are: *nišmarti, nišmartem* and *nišm'ru.*

2.2.3 Piel

The *piel* is often an intensive of the *kal,* but equally often indicates no special meaning. It is typified by two bimoraic syllables in the base, usually surfacing as CiCCeC.[15] While most approaches assume that the vowels /i/ and /e/ are part of the specification of the paradigm, I assume here that only the syllable structure and the vowel /e/ are specified, with the /i/ coming from epenthesis (as described below).

From d.b.r we find *diber* 'he spoke.'[16] From k.r.s.m we find *kirsem* 'he ravaged.'

With pronominal endings we find *dibarti* 'I spoke,' *dibartem* 'you (pl.) spoke' and *dibru* 'they spoke.'

2.3 Biliteral Roots

One of the major triumphs of McCarthy (1981) was the elegance with which it handled spreading of consonantal radicals in Arabic. In Hebrew, however, the situation is considerably more complex. It is one of the main goals of the present work to explain the surface forms, so for now they are simply given, that the reader might know what is to be explained.

The examples use the root s.b 'spin.'

2.3.1 Kal

In TH, we find both *sav* and *savav* 'spun.' With affixes, we find: *saboti* 'I spun' (instead of the *savti* or *savavti* one might expect), *sabotem* 'you (pl.) spun' and again both *sav'vu* and *sabu* for "they spun."

In MH, we find *sav* and *savav*; *savti*; *savtem*; and *savu* and *sav'vu.*

2.3.2 Niphal

In TH, we find *nasav* 'was spun,' instead of the expected *nisbav.* The other forms are *n'saboti, nasabu,* and *n'sabotem.* (Cf. the unattested: **nasavti* or **nisbavti*, etc.)

[15] Often, however, the medial V's will be identical, and hence geminate, and then undergo degemination.

[16] The /e/ in this form is unexpected. In TH, it alternates with /a/, as in *dibar.* In both TH and MH, all other forms (those with suffixes) have /a/.

In MH, we find *nasav, nasavti, nasavu* and *nasavtem.*

2.3.3 Piel

In both TH and MH we find *sovev* 'spun,' instead of the expected *sibev.* The other forms are: *sovavti, sovavtem*[17] and *sov'vu.*

2.4 "Full" Biliteral Roots

Throughout most of this article, by "biliteral root" will be meant "biliteral roots that do not behave as triliteral roots," as in the discussion above. However, many biliteral roots behave exactly as though they were triliteral roots with their second and third radicals identical. At least two roots behave as though their first and second radicals are identical. Due to space constraints, these will be paddressed here only in passing,[18] but the general approach to be taken regarding these roots is that the OCP does not hold of every Hebrew root, and some roots contain identical adjacent consonants.

3 Constraints and Paradigms

3.1 Bases

I assume here that each final verbal form in Hebrew contains a base and optionally affixes. Each paradigm, then, is defined as a hierarchy of constraints on the base and a hierarchy of constraints on the affixes. The constraints on the final form of the verb do not depend on the paradigm, and in addition to the paradigm-specific constraints, language-wide constraints apply to the bases as well.

The simplest case is the *kal* paradigm, which specifies no affixes and but one constraint on the base, namely, the last syllable of the base be bimoraic.

For regular verbs, as we saw above, three root letters, say, q.t.l, form a base of the sort "[qa.tal]."[19] To this base various affixes are added.

In the case of biliteral verbs, say, q.l, the base will be "[qal]." Notice that [qatal] and [qal] have in common only the vowel /a/ and the fact that both have a bimoraic base-final syllable. The base [qatal] cannot be fulfilling the requirement that "a base be a foot" because that does not obtain with "[qal]." Clearly, "[qal]" cannot be the result of putting q.l into a "CaCaC" paradigm.

Niphal behaves exactly as *kal,* with only one constraint on the base, but with the additional requirement of a prefixed /n/. So, the *niphal* form from q.t.l, *niqtal,* is formed from a base "[qtal]"

[17] Or perhaps *s'vavtem* in TH — data are lacking.
[18] Some examples are shown in Table 1.
[19] I will use square brackets to delineate bases.

to which *n* is prefixed. (See the discussion below for why [qatal] is optimal in *kal* but [qtal] optimal in *niphal.*)

Piel differs from both *kal* and *niphal* in that it specifies two bimoraic syllables for the base. So, still from q.t.l, we find a base of "[qit.tal]." The biliteral verbs, such as s.b, pattern thus: "[saa.vev]." (Again, see the text below on why each is optimal.) *Piel* has no specified affixes.

The *hitpael* paradigm (not discussed below) behaves exactly like *piel* except that it specifies prefixation of *hit-* to the base.

3.1.1 Vowels

With one exception, every vowel of the three paradigms examined here can be predicted without referring to the paradigm, and so it seems a reasonable step to assume that paradigms are not specified for vowels at all. Some vowels, as we will see below, must not be specified at part of the paradigm, for they change according to other factors.

The final vowel in every paradigm is /a/. From this, we might assume that /a/ is a default vowel, and so that the paradigmatic[20] requirement of bimoraic syllables is met by inserting an /a/. However, this cannot be an epenthetic vowel, for /i/ is the epenthetic vowel in Hebrew. Further, there is one form, the 3-m-sng *piel* form, that has an /e/ instead of the expected /a/. The passive paradigms, not discussed here, seem to have /u/ specified as their first vowel. What we have is the not untypical case where the nicest generalization seems to be just that: a generalization. Fortunately, for our purposes here, it does not matter where the /a/ in the roots comes from. Everything here is compatible both with the hypothesis that the /a/ is part of the specification of the paradigm and with the hypothesis that the /a/ is inserted as the default vowel.

By contrast, the explanation here crucially depends on the fact that the first vowel in some of these paradigms not be specified, contra McCarthy (1981) and Bat-El (1989).

3.2 On "Bimoraic" Syllables

Hebrew seems to be unique in that the moraicity of syllables is used as a formative. Further, bimoraic syllables seem to be prohibited anywhere they are not needed.[21] So, for example, returning to *kal,* we find that the base from q.t.l must be [qatal], and not, say, [aqtal], or [qaa.tal] because the latter two contain two bimoraic syllables, but only one is required. We cannot assume, however, that independent foot constraints prohibit two consecutive bimoraic syllables, because *piel* is defined precisely as having two consecutive bimoraic syllables.

[20]Alas, the word "paradigmatic" is now ambiguous, having its usual sense meaning, roughly, "normal" or "typical," and its technical sense here of "of or relating to a paradigm."

[21]The prohibition against spurious bimoraic syllables accords nicely with their use as formatives. In general, formative material cannot appear spuriously.

So I assume here that bimoraicity is possible only when needed. One nice way of capturing this is a low-ranked constraint MonoMora, which dictates that, other things being equal, monomoraic syllables are preferred over bimoraic ones. However, for simplicity herein, I will not even consider spurious violations of MonoMora, that is, bimoraic syllables in a base not resulting from paradigmatic requirement for such.

It is not clear how best to encode the bimoraic requirements in terms of constraints needed elsewhere. One possibility, of course, is that Hebrew is unique not only in the way it makes use of syllable weight but also in the way it specifies that syllable weight. But we should prefer a more systematic solution. Perhaps this might be best accomplished with alignment constraints, though the simplest solution will clearly not work.[22] Whatever the case may be, we may forge on, using the descriptive result that the paradigms dictate syllable weight, and hoping that future research will place the pattern in a larger framework.

3.3 Final Forms

The final form of a verb in a given paradigm is thus built as follows. Bases are constructed, by Gen, from the root consonants according to the constraints of the paradigm. To those bases affixes are added, again according to the constraints of the paradigm; this presumably is also the job of Gen, but as the affixation constraints appear to be unviolated, Gen does little work here.

Along with the paradigm-specific affixes, agreement affixes may be added to the base. In most cases, several bases will meet the constraints of the paradigm, and so several combinations of affixes and bases will be given to Gen, from which the final form of the verb is created.

Because Gen is applied twice, first to form the base and then to form the final verb form, even the "inviolable" constraints that are used to form the base may be violated, by the second application of Gen.

Notice, however, that the soft constraints are only applied once. The constraints in the morphology, those that create the stem, do not simply create the most optimal stem and then hand it to the phonology. Rather, as we will see below, the stem that will create the best output for any given stem-affix combination is the stem that will be used.

4 The Constraints

I assume the following constraints:

[22] If we assume that bimoraic final syllabicity is encoded as an alignment constraint that the right edge of a bimoraic syllable be aligned with the right edge of base, we will account for *qal* and *niphal,* but the equivalent constraint for *piel,* namely, that the left edge of a bimoraic syllable be aligned with the left edge of a base, will incorrectly predict that one syllable might suffice to satisfy both constraints at once.

NoCoda Syllables cannot end in a consonant.

Parse Material in the input must be in the output. We will quickly see that this must be parameterized, so that changing /b/ into /v/ is less of a violation than changing /b/ into /r/. One way to do this is to have the Parse constraint operate at the feature level. But one might still wonder why /b/ turning into /v/ is better than /b/ turning into /p/, especially as that latter case is attested in some voice assimilation (which will not be discussed here). Of course, we might might create a different parse constraint for each feature: Parse_{voice}, Parse_{cont} etc. But to do so would be to fail to recognize that the various parse constraints are related one to another. That is, Parse_{voice} and Parse_{cont} are related in a way that other constraints are not. However, the exact nature of Parse is far too daunting a topic for us now, and so we must at times rely on our intuition that a /v/ in the final form from a /b/ in the input form is less of a violation of Parse than an /r/ or /p/ in the final form would be. The details of this implementation, while crucial for OT at large, will not affect the results here.

So for our purposes here, we assume at least two constraints, $\text{Parse}_{R(andom)}$ and Parse. The former is violated when the "wrong" element is chosen by Gen not to be parsed, as in the b/r case above. Inserting /r/ for /b/ would violate Parse_R, while inserting /v/ would only violate Parse. We will quickly see that Parse_R is unviolated. Likewise, not parsing half of a geminate cluster violates Parse but not Parse_R.

A much more systematic treatment of these issues is possible, but prohibited here by space constraints.

Fill Material in the output must come from the input. Again, as before, we must recognize that certain violations of Fill are worse than others. For one thing, where epenthesis is required even though it violates Fill, only one epenthetic vowel is possible, /i/.

A second issue involves the fact that epenthesizing a vowel in the right circumstance, while a violation of Fill, must be a violation of a sort different that inserting a random sound somewhere. To capture this, we need, as with Parse, two constraints: Fill_R and Fill, only the latter of which is violated by epenthetic vowels.

NoAA The long vowel /aa/ does not exist in Hebrew. In its stead we find /o/.[23] By ranking NoAA highly, we account for this, again assuming that Fill and Parse behave appropriately.

[23] The evidence that /aa/ surfaces as /o/ — though supported by data from other languages — is largely theory internal, unfortunately.

NoStop The consonants /bgdkpt/ (/bkp/ in MH) spirantize after vowels and at the end of a word. We can accomplish this with a constraint against the sequence VC_1C_2, where C_1 is the appropriate consonantal stop, and C_2 is a different consonant, or — as is often the case — the end of the word. Certainly we should like a more precise formulation of this, but the results should be clear.

Fortunately, many languages exhibit post-vocalic spirantization, and so we have cross-linguistic evidence for NoStop.

NoGem In spite of claims to the contrary, Hebrew does not permit geminates on the surface. We can know this about MH simply by listening to the language.[24] The evidence that TH had no geminates is more complex, but the reader will do well to note that the evidence in the literature about geminates in TH is largely theory internal, and was used to make spirantization work properly. Also, comparison with cognate languages, such as Arabic, suggests geminates. But there is little evidence that the surface forms had geminates, and, as we shall see, assuming that, like MH, TH did not allow them helps us understand the verbal system more completely.

$\text{Align}_{\textit{rt-ft-L}}$ **and** $\text{Align}_{\textit{rt-ft-R}}$ Following Sharvit (1994), I assume that roots (her terminology) or bases should be, when possible, left aligned and right aligned with feet. These two constraints capture that.

Stress This is not really a constraint, but will be used to mark forms that violate Hebrew stress patterns. Those patterns are far too complex to discuss here, but for a few forms we will need to make use of them. When one form is optimal because a related one violates the stress pattern of Hebrew, the second form will be marked *Stress. However, Stress is just a useful abbreviation, and not a constraint, and so it cannot be ranked.

5 Constraints and Verbs

5.1 Full Roots in TH and MH

5.1.1 Kal

Our first task is to understand how full roots acquire the form they do, and we begin by considering the base form in the first paradigm (*pa'al* or *kal*). From the root q.t.l we need to derive *qatal*. As indicated above, I assume that this paradigm is marked by the vowel /a/ and the syllable requirement that the final base syllable be bimoraic.

[24] Actually, the matter is more complex, because there are some geminate clusters, but not where they are traditionally described to be. When the 1-pl suffix *-nu* is added to roots ending in /n/, we sometimes find gemination.

There are two stages here, first morphology, then phonology. The morphology takes the root consonants and the vowel and creates appropriate bases. So, *qatal, aqatal, qtal, qatalal, qatalaqatal,* etc. are all possible bases. Remember that a base syllable can only by bimoraic if is marked by the paradigm to be bimoraic, and so *qatqatal,* for example, is not a possible *kal* base. We also need not even consider a base *[qatala],* which happens to be the equivalent Classical Arabic form, because it violates the morphological requirements of the base, and morphology is assumed to dominate phonology. Of course, we cannot use this reasoning to rule out the final form *qatala,* only such a base. That is, we might still have *[qata.l]a.*

This raises an important point. In the final form *[qa.ta.l]a,* the base no longer ends in a bimoraic syllable. But it does not violate the paradigm's constraints, because, at the point at which those constraints applied, the base did conform.

From the above, we have at least the following:

(1)

		Parse$_R$	Fill$_R$	Align	NoCoda
(a)	[<q>tal]	*			*
(b)	[qital]		*		*
(c)	[qatalaqatal]			*	*
(d)	[qatalal]			*	*
(e)	[qata.l]a		*		
(f)	[aqatal]			*	*
(g)	[aq<a>tal]	*		*	*
(h)	☞[qatal]				*

From these we see immediately that Fill$_R \gg$NoCoda. The form *qatal* obtains in both TH and MH.

The suffixation of *-ti* 'I' is straightforward: *qatál+ti.* That form, like the base, violates only NoCoda.

The suffixation of *-u* 'they' is more interesting:

(2) a. *[qatal]ú   *Stress

b. [qat.l]ú   *NoCoda

The expected form, *qatalú* actually loses to *qatlú.* That this is a consequence of stress assignment can be seen from TH, in which phrase final stress is sometimes shifted back one syllable. In those cases, only *qatálu* is possible. Stress-related vowel shortening is beyond the scope of this article.

One other suffix is relevant, *-tem* 'you (pl.).' We find:

(3) a. (MH) [qatáltem] **NoCoda

b. (TH) [q'taltém] **NoCoda

both of which doubly violate NoCoda. This pair further demonstrates the role of stress in vowel shortening.

5.1.2 Niphal

We consider next the *niphal* paradigm, marked, like *kal,* for a bimoraic final syllable, but also for a prefixed *n-*. The options for creating a base are identical to the *kal,* but the final forms are different because of the prefix.

(4)

	Parse_R	Fill_R	Align	NoCoda	Fill
(a) na[q.tal]		*		**	
(b) na[qa.tal]		*		*	
(c) n[a.qa.tal]			*	*	
(d) <n>[qatal]	*			*	
(e) n[<q>atal]	*			*	
(f) ni[qatal]		*			
(g) ☞ni[q.tal]				**	*

The only difference between (4a) and (4g) is that the former contains the wrong epenthetic vowel. This is reflected in the different constraints violated by each, Fill vs. Fill_R

My assumption that the /i/ in the *niphal* is epenthetic varies from standard accounts (which assume that it is part of the specification of *niphal*).Here it is mildly more general to assume that the vowel comes from epenthesis, but below we will see that the correct weak verbal forms can be derived only with this assumption. There is also considerable evidence outside the verbal paradigm to indicate that an epenthetic /i/ is inserted after the first of three consonants at the beginning of a word.[25]

Suffixation to *niqtal* proceeds as expected based on the *kal* paradigm.

[25]For example, there is a class of clitic prefixes in Hebrew, among them *l-* 'to' and *b-* 'in,' that normally are prefixed without a vowel, or with a shewa, between them and another word. However, before a word beginning with two consonants, an epenthetic /i/ is inserted.

5.1.3 Piel

Old Roots. The next paradigm is *piel,* marked by two bimoraic syllables. Here we see a different set of possible bases, of which *qitlel,* and *qittel*[26,27] are reasonable, assuming, as before that the /i/ is inserted epenthetically.

(5)

	NoGem	NoCoda	Fill	Parse
(a) [qit.lel]		**		
(b) [qit.tel]	*	**		
(c) ☞[qi<t>.tel]		*		*

The data in (5) bring out an important aspect of Hebrew, and of the dual-level-nature of the system I describe here. We must assume that Hebrew has a surface constraint against geminates, ranked very high. Certainly MH only rarely exhibits (overt) geminates, and there is only theory-internal evidence to suggest that TH has them. But pre-surface geminates play a crucial role in spirantization, a topic I discuss below.

Notice that the final base form in (5c) no longer has two bimoraic syllables. This is unproblematic, however, because there is no restriction against mucking around with a base once it is formed.

Because *qitel* wins out over *qitlel*, we see that NoCoda≫Parse.

New Roots. The situation in MH in *piel* is considerably more complex. While old roots (from TH and even later) continue to appear only as (5c) and not (5a), modern roots appear both ways. Much attention has been given to the MH form *fikses* 'he faxed,' (from the English /faks/) wherein the final /s/ spreads, rather than the medial /k/ doubling. This appears to be evidence that, in MH, Parse≫NoCoda. However, along with *fikses* we also find *fikes* 'he focused,' from the English "focus." In this latter case, NoCoda once again wins out. Two conclusions are possible.[28] Either MH does not rank these two constraints, or there are two systems at work in MH, each with a different ranking. Because the vast majority of new roots follow the *fikses* pattern, and because we will need this conclusion below, I assume that MH ranks Parse over NoCoda, and that the (MH) word *fikes* 'focused' is a remnant of the older system.

[26] The /e/ as the final vowel is difficult to explain. In TH, it alternates with (the expected) /a/, but in MH it does not. Further, the /e/ appears only in the third person masculine, /a/ appearing in its stead for the other conjugations in both TH and MH. We must assume this /e/ is simply a lexical quirk.

[27] Or *qtlel* and *qttel.*

[28] Actually, a third possibililty also presents itself: One word is from the root f.k.s and the other from f.k.s.s. In fact, I suspect this is right, but space constraints prohibit a full discussion of this possibility.

This is the first example we see of a way in which MH declension differs from TH. In TH, NoCoda≫Parse. In MH, the reverse is true.

5.2 Spirantization and Gemination

One final refinement in our treatment of full roots in necessary. Recall that some sounds become spirantized after a vowel in Hebrew, except when geminate. In TH, these sounds are /bgdkpt/; additionally, TH does not have phonemic spirantized versions of these sounds. In MH, on the other hand, only /bkp/ undergo spirantization. To make matters more complicated, MH *does* have spirantized phonemic equivalents of all three sounds. Ideolectal variation often determines whether a TH form in which a spirantized sound appears is incorporated into MH as a spirantized sound or a phonemic sound.

At any rate, it is clear that spirantization must occur "before" degemination, so that from *dibber* we get *diber* and not *diver*. Recall that both processes are assumed here to be the result of constraints (NoGem and NoStop) ranked higher than Parse. NoStop only applies to certain configurations. If degemination (e.g., "<t>t") is not one of the configurations to which NoStop applies, we will achieve the result that spirantization occurs "before" degemination.

A few examples with the roots š.b.r and k.p.c will help:

(6) a. *kal:* šavar, kafac

b. *niphal:* nišbar, nikfac (MH) or nikpac (TH).

c. *piel:* ši<b>.ber, ki<p>.pec

Notice in particular (6c), in which the medial /b/ and /p/ remain because of their earlier geminate status. In (6b) we see phonemic /f/ in MH.[29]

Notice too that the TH forms which have undergone spirantization violate Parse, while the MH forms do not (because the spirantized form is underlying in MH). We now have a second way in which TH and MH differ. In the former NoStop≫Parse, while in MH the opposite is true.[30]

5.3 Biliteral Roots in TH

Because biliteral behave so differently in TH and MH, it will be convenient to discuss them separately. We consider TH first.

[29] Actually, *nikfac* isn't much a word. A much clearer example comes from infinitives, *likpoc* in TH and *likfoc* in MH.

[30] NoStop is probably inactive in MH.

5.3.1 Kal

Consider now a root such as s.b 'spin,' which is generally assumed to consist of only two letters.[31] How might this root appear in the *kal?*

As before, the main requirement of *kal* is that the final syllable of the base be bimoraic. Among the sensible bases, then, are *sav* and *savav,* the others being undesirable for reasons discussed above. In fact, both are attested:[32]

(7) a. [savav] *NoCoda *Fill[33]

b. [sav] *NoCoda *Fill

In TH, the letters that can do so, spirantize word-finally, regardless of the their geminate status.[34] That is why we see /v/ and not /b/ in (7). In order to allow both possibilities, we must assume that both instances of /v/ in (7) count as the same Fill violation.

Next, we add the suffix *-u* 'they.' Here we find a minor surprise. Again, two forms are possible:

(8) a. [sav'v]ú (or, in pause, [saváv]u)

b. [sáb]u

Why doesn't the /b/ in (8b) spirantize? The only possible answer can be that it was once geminate, which means that we have misunderstood the base form. Specifically, consider the following:

(9)

	NoGem	NoStop	NoCoda	Fill	Parse
(a) [sab'b]u		*	*		
(b) [sav'v]u			*	*	
(c) [sa.v]u				*	
(d) ☞[sa<b>.b]u					*

From this we see that Fill≫Parse. What we have not accounted for is the well-formedness of (9b) (cf. the ill-formed (9c)). In general, this is a good thing, as most roots do not allow the option in (9b). The most likely solution to this apparent dilemma follows from my hypothesis that both

[31] Traditional Hebrew grammars consider these roots to be triliteral, with a doubled last radical. This approach may actually be better, for reasons that are too detailed to discuss here.

[32] Sadly, they do not alternate freely. The shorter form is intransitive, and the longer one transitive, at least in TH.

[33] Because of the b/v alternation

[34] With one minor exception which need not concern us.

s.b and s.b.b are possible roots. While much evidence exists in support of this, a full discussion of it is too lengthy for the present context.

The really interesting case concerns the affixation of consonant-initial suffixes, such as *-ti* 'I' in TH. Rather than **savavti* or **savti,* we find only the form *saboti.* This is because of NoCoda, as we see immediately below. To understand the forms, we assume that long /a/ surfaces as /o/, as discussed above. The constraint NoAA takes care of this.

(10)

		NoGem	NoStop	NoCoda	Fill	Parse
(a)	[savav]+ti			*	*	
(b)	[sav]+ti			*	*	
(c)	[sa<b>b]+ti			*		*
(d)	[savo]+ti				**	
(e)	☞[sa<b>bo]+ti				*	*

In (10d) we find one Fill violation for the /o/, and one for the /v/. By contrast, in (10e) we have only violation, for the /o/.

The pattern seen in (10) is further confirmation that Fill≫Parse.

The good form comes from the base [sa.bbaa]. Because of the biliteral nature of the root, we can double the vowel to form the required bimoraic syllable in the base, and so avoid the NoCoda violation.

This is a striking validation of the original assumption that the paradigms are defined by syllable weight and not by C-V templates. If the *kal* paradigm were indeed defined simply by CVCVC, as in McCarthy (1981), we should have no way of understanding the spreading of the V to a C slot. If we assume that the final form is dictated only by syllable constraints on the word, as in Bat-El (1989), we have nothing to contra-indicate **savati.* It is only by assuming that the final syllable of the base must be bimoraic that we can correctly predict the base in these forms.

These examples also demonstrate again the importance of NoCoda in TH.

5.3.2 Niphal

We consider next what happens with with these biliteral roots in the *niphal* paradigm.

Recall that in the case of the triliteral roots we found that the best base was [qtal], and that an epenthetic /i/ would be inserted later in the derivation. Other approaches have assumed that the /i/ is part of the paradigm. Here we will see the benefits of the current system.

Above, [qtal] was chosen over [aqtal] because that latter form would create a bimoraic first syllable, and we have assumed that bimoraic syllables in roots can only be the result of paradigmatic specifications. Here, we must consider all of [sbab], [sab] and [asab]:

(11)

	FILL$_R$	NOCODA	FILL
(a) ni[s.bav]		**	*
(b) ni[sav]	*		*
(c) ☞n[asav]		*	*

Indeed, *nasav* is the correct form.

Other accounts, which assume that *niphal* is defined by (*inter alia*) the vowel sequence "i.e" have no way of explaining (11).

The suffixed forms of this paradigm behave as *kal,* yielding *n'sabóti* and *nasábu.* Again the stress-related vowel reductions are beyond the scope of what can reasonably be covered here.

5.3.3 Piel

We turn finally to the *piel* paradigm, wherein the requirement for the base is that it consist of two bimoraic syllables. From the root s.b we have at least three reasonable ways of doing this: [sb.beb] (as with the triliteral roots), [saabeb], or [sibseb]:[35,36]

(12)

	NOCODA	FILL	PARSE
(a) [sivsev]	**	*	
(b) [si<b>.bev]	*	**	*
(c) ☞[so.vev]	*	**	

The first form is rare in TH. The second two general forms are attested, but usually not with the same verb. In this case, we find *sovev* and not **sibev.* But we also find *hilel* 'praised' and not *holel.* This is another instance of extra-phonological determination of form. Most likely, the "regular" form *hilel* comes from a root h.l.l.

Here, too, though, we see clear confirmation that the paradigms dictate syllable structure only, and that the /i/ of the *piel* form is not part of the paradigm's specification (contra McCarthy (1981) and Bat-El (1989)).

[35] Remember that the final /a/ becomes an /e/ for unknown reasons.

[36] Space constraints prohibit a discussion of the reduplicated pattern demonstrated by *sibseb,* which, while not a possible form for the root s.b, is attested with other roots.

5.4 Biliteral Roots in MH

Remember that MH has a significantly lowered ranking of NoCoda w.r.t. Fill and Parse. That ought to change matters considerably with regard to biliteral roots, and it does.

5.4.1 Kal

We consider the two forms from (7) again, now in the context of MH:

(13) a. [savav] *NoCoda

b. [sav] *NoCoda

As in TH, both are possible. Notice too that there are no Fill violations because the /v/ is now underlying.

With the suffix *-u,* we find, as expected, some confusion regarding the b/v alternation (because most but not all speakers have adopted a morphemic /v/ in this root), and so three forms are possible:

(14) a. [sav'vú]

b. *[sábu] *Fill$_R$[37]

c. [sávu]

Example (14b) is good only for those speakers for whom the /v/ is not morphemic. Such speakers are rare.

The real difference comes with the consonant-initial suffixes, such as *-ti.* Recall from above that *saboti* was chosen over both *savti* and *savavti* because those two forms contained a NoCoda violation. In MH, Parse is ranked higher than NoCoda:

(15)

	Fill	Parse	NoCoda
(a) [savo]+ti	*		
(b) [sa<b>b]+ti		*	*
(c) [sa<b>bo]+ti	*	*	
(d) ☞[savav]+ti			*
(e) ☞[sav]+ti			*

[37] Some speakers still have an underlying /b/ here, and this form is okay for them.

In (15c) the Fill violation results from the /o/. In (15d) the Parse violation results from the <b>.[38]

Here, in complete contrast to TH, the forms with the NoCoda violations are chosen over those with the Parse or Fill violations.

5.4.2 Niphal

Recall from (11) that the correct form, *nasav,* depends only on the number of NoCoda violations. Accordingly, we correctly expect that forms identical to TH will obtain in MH:

(16)

		Fill$_R$	NoCoda
(a)	ni[s.bav]		**
(b)	ni[s.vav]		**
(c)	ni[sav]	*	
(d)	☞n[asav]		*

The derived forms in MH follow the same pattern seen immediately above with the *qal: nasavu* and *nasavti.* The first of those wins out against *nasvu* because it contains one fewer NoCoda violation; in the second case, *nasavati* is impossible because the base would not contain a final bimoraic syllable, and any form that ends *-oti* is impossible for the same reasons it was in *qal,* namely, it would constitute a Fill violation.

5.4.3 Piel

As with TH, we expect three possible forms, *sivsev, sovav* and *sibev*. Unlike TH, however, where only the latter two are widely attested (but with different roots) in MH we find all three.

The personal affixes do not change the base form.

5.5 Ordering of Constraints

Here's what we know about the constraints in Hebrew.

(17) Constraint ranking in TH:

a. Fill$_R$ and Parse$_R$ are unviolated.

b. NoAA is unviolated.

[38] This may also be a violation of Fill$_R$.

c. NoCoda≫Parse

d. Fill≫Parse

e. NoGem≫Parse

f. NoStop≫Parse

(18) Constraint ranking in MH:

a. Fill_R and Parse_R are unviolated.

b. NoAA is unviolated.

c. Parse≫NoCoda

d. Fill≫Parse

e. (From the above we deduce that Fill≫NoCoda.)

f. Parse≫NoStop

The crucial difference between TH and MH is that in the latter a greater importance has been placed on Parse. That minor reordering produced massive changes in the declension of the weak verbs. In the next section, I discuss why a greater importance should have been assigned to those constraints in MH.

6 How TH Became MH

So far we have achieved considerable success. By re-ranking one constraint, we have accounted for a systematic change in behavior between TH and MH. However, we can go one step further, and in fact begin to understand *why* Parse achieved greater importance in MH. Space considerations prevent a detailed analysis, but the basic idea is as follows.

The greater importance assigned to Parse in MH reflects a requirement that final forms more closely resemble underlying forms. In TH, the most prevalent fidelity violations stemmed from the spirantizaion of /bgdkpt/. But in MH, three factors have destroyed the regular rule of spirantizaion, and so the wide-spread fidelity violations can be attributed to other factors, markedly decreasing the evidence for Parse violations.

The three factors, already alluded to above, follow

First, the coalescence of the pairs /qk/, /xχ/ and /vw/. The first forced a modification of the rule that would turn /k/ into /x/. Now only some /k/'s, those that were not originally /q/, turn in /x/. The second two worked the other way around, creating a phonemic /v/ not derived from /b/ and a phonemic /x/ not derived from /k/.

The second factor is the disappearance of spirantized versions of /tdg/ (that is, /θðγ/) which further weakened the once-regular rule of spirantization.

Finally, loan words introduced phonemic /fvx/.

Because of these three factors, the systematic rule of spirantization seen in TH has become erratic, and forms that once could only be analyzed as involving PARSE violations (that is, spirantization) might now represent underlying phonemic inventory. But the main stimuli for ranking PARSE so low come from spirantization. To properly rank NOSTOP, learners must know that TH has no phonemic spirantized versions of the stops. Likewise, to learn that NOGEM dominates PARSE, they make use of the fact that spirantization fails to occur where it ought to, which in turn depends on knowing that spirantization ought to have occurred were it not for the (pre-surface-form) gemination. The ranking of PARSE below NOCODA does not depend directly on spirantization, but depends indirectly on the existence of other constraints outranking PARSE.

So the change in underlying inventory coupled with the change in spirantization caused a re-ranking of PARSE, which in turn caused a massive restructuring of biliteral-root forms.

7 Summary

We have seen that previous accounts of Hebrew morpho-phonlogy do not adequately address biliteral roots in Hebrew or the change in verbal paradigms seen between TH and MH. The current account, working within an OT framework, accounts for this expanded set of data by using moraicity requirements as formatives in the verbal paradigms and assuming that vowels hitherto thought to be part of C-V templates in fact derive from more fundamental concerns. The change from TH to MH is captured in terms of the re-ranking of one OT constraint, PARSE, and possible factors (external to the data to be explained) for that reranking were presented. This pattern helps validate both the details of the theory presented here and the larger theory of OT.

References

Bat-El, O., 1989. *Phonology and Word Structure in Modern Hebrew*. UCLA dissertation.

McCarthy, J. 1981. A prosodic theory of nonconcatenative morhpology. *LI* 12.373–263.

Prince, A., & P. Smolensky. to appear. *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge, MA: MIT Press. Also avail. as ms. 1993 from Rutgers University and University of CO.

Sharvit, Y., 1994. Issues in the phonology and morphology of the Modern Hebrew verbal system. ms. Rugters University.