# Consonant Identity and Consonant Copy: The Segmental and Prosodic Structure of Hebrew Reduplication

*Outi Bat-El*

The article addresses two issues regarding Hebrew reduplication: (a) the distinction between reduplicated and nonreduplicated stems with identical consonants (e.g., *minen* 'to apportion' vs. *mimen* 'to finance'), and (b) the patterns of reduplication ($C_1VC_2VC_{2C}$, $C_1VC_2C_3VC_{3C}$, $C_1VC_2C_{1C}VC_{2C}$, and $C_1C_2VC_3C_{2C}VC_{3C}$). These issues are studied from a surface point of view, accounting for speakers' capacity to parse forms with identical consonants regardless of their base. It is argued that the grammar constructed by the learner on the basis of structural relations (base – output) can also serve for parsing surface forms without reference to a base.

*Keywords:* Hebrew, reduplication, prosodic structure, segmental identity, morphological parsing, Optimality Theory

Words in Modern Hebrew (hereafter Hebrew) may have one or two pairs of identical consonants anywhere in the stem. I argue that the stem is reduplicated when the pair appears (or the pairs appear) at the right periphery, without an intervening consonant (e.g., *garar* 'to drag', *ʃixrer* 'to release', *bakbuk* 'bottle', *ʃravrav* 'plumber'); otherwise, the stem is not reduplicated (e.g., *mimen* 'to finance', *diskes* 'to discuss', *safsal* 'bench', *ʃóreʃ* 'root'). The reduplicated stems have three possible prosodic structures, CVCVC, CVCCVC, and CCVCCVC, which can host one or two pairs of identical consonants. Consequently, there are four patterns of reduplication: $C_1VC_2VC_{2C}$, $C_1VC_2C_3VC_{3C}$, $C_1VC_2C_{1C}VC_{2C}$, and $C_1C_2VC_3C_{2C}VC_{3C}$ (where subscript *c* marks the copy).

In this article, I explore Hebrew stems with identical consonants, with emphasis on their surface structure (i.e., without reference to a lexical base), rather than their derivation. I address the following questions:

> How do speakers assign structure to forms with identical consonants, distinguishing between reduplicated and nonreduplicated stems?
> What is the system responsible for the surface patterns of reduplication?

I acknowledge the contribution of those who provided comments on and suggestions about earlier drafts of this article: Adam Ussishkin, David Gil, Iris Berent, and my students in the Linguistics Department at Tel Aviv University; the participants in the colloquia at University of California at Santa Cruz, Bar Ilan University, and Tel Aviv University; and two anonymous reviewers whose comments were remarkably insightful. The usual disclaimers apply.

Linguistic Inquiry, Volume 37, Number 2, Spring 2006
179–210

Studies of Hebrew reduplication have mostly been concerned with the relation between a lexical base (either an abstract underlying representation or a surface form) and its reduplicated counterpart. The empirical basis of these studies has been denominative verbs (McCarthy 1979, 1981, 1984, Bat-El 1989, 1994b, 1995, Gafos 1998, Ussishkin 1999, 2000) and the nominal/adjectival pattern $C_1C_2VC_3C_{2C}VC_{3C}$ (McCarthy 1979, 1981, Graf 2002).

There are several reasons to also study reduplicated forms without reference to their lexical base: (a) there are nonreduplicated forms with identical consonants in different positions in the stem (e.g., *mimen* 'to finance'); (b) there are many ''orphan'' reduplicated forms—that is, forms without a nonreduplicated counterpart and thus without a surface lexical base (e.g., *garar* 'to drag'); (c) experimental studies (see section 2) suggest that speakers distinguish between reduplicated and nonreduplicated nonce words with identical consonants. A derivational account can assume an abstract underlying representation for the orphan reduplicated forms, but not for the nonce words. It should be clarified that I do not say that there is no derivational process of reduplication, given speakers' capacity to form a new reduplicated form from a given base. Rather, I claim that speakers can identify a reduplicated form without reference to a lexical base, and it is this capacity I wish to account for here.

Many of the studies mentioned above are also concerned with the patterns of reduplication. Those that study all the patterns provide a different analysis for each one. Here, I provide a unified account for all patterns of reduplication, arguing that there is no independent reason for different analyses.

I start the discussion with the relevant data, showing that (a) in terms of prosodic structure, vocalic pattern, and affixation, reduplicated forms are identical to nonreduplicated ones, and (b) the prosodic restrictions (the minimal word) are not sufficient to account for the triggering of all cases of reduplication (section 1). I proceed with the problem that early Optimality Theory analyses of the $C_1VC_2VC_{2C}$ pattern encounter in light of nonreduplicated forms with identical consonants, namely, $C_{i1}VC_{i2}VC_{j3}$ (section 2). I solve the problem by providing a set of ranked constraints that allows speakers to parse forms with identical consonants and distinguish between reduplicated and nonreduplicated forms (section 3). Then I provide a unified account for all the patterns of reduplication, proposing that the morphological constraint triggering reduplication has three different positions within a fixed ranking of prosodic constraints (section 4). The unified approach is further supported by the absence of a coherent semantic distinction among the different patterns of reduplication (section 5). The conclusion also highlights the advantage of a reduplication constraint over a reduplication affix (section 6).

## 1 Relevant Language Background

Hebrew words may consist of a bare stem or a stem plus affixes. Derivational relations are expressed by apophony (i.e., alternation in vocalic pattern, ablaut), affixation, and alternation in prosodic structure, the last often accompanied by affixation.[1]

[1] Throughout the article, affixes are underlined and stress is final unless otherwise specified.

(1) *Derivational relations among Hebrew words*

| Verbs | | Nouns/Adjectives | |
|---|---|---|---|
| gadal | higdil | gadol | migdal |
| 'to grow' | 'to enlarge' | 'big' | 'tower' |
| daxaf | nidxaf | daxuf | madxef |
| 'to push' | 'to be hastened' | 'urgent' | 'propeller' |
| xideʃ | hitxadeʃ | xadaʃ | xadʃan |
| 'to renew' | 'to be renewed' | 'new' | 'innovator' |
| halax | hithalex | halix | halixon |
| 'to walk' | 'to walk about' | 'process' | 'treadmill' |
| hixlif | hitxalef | xalif | taxlif |
| 'to change' (trans.) | 'to change' (intrans.) | 'exchangeable' | 'substitute' |

With the exception of forms of the pattern $C_1C_2VC_3C_{2C}VC_{3C}$ (see section 5), reduplicated words are structurally identical to nonreduplicated ones, in terms of vocalic pattern, prosodic structure, and affixation (hereafter ''configuration''). That is, a configuration can host either reduplicated or nonreduplicated forms. This is shown in (2) with a variety of verbal and nominal configurations.

(2) *The surface-structural identity of reduplicated and nonreduplicated forms*

| | Configuration | Reduplicated | Nonreduplicated |
|---|---|---|---|
| Verbs | | | |
| | CaCaC | gazaz | gazam |
| | | 'to chop' | 'to cut' |
| | CiCeC | cided | gidel |
| | | 'to side with' | 'to raise' |
| | CiCCeC | ʔiʃrer | diklem |
| | | 'to reconfirm' | 'to recite' |
| | CiCCeC | kilkel | tilfen |
| | | 'to spoil' | 'to phone' |
| | hiCCiC | hiflil | hixnis |
| | | 'to incriminate' | 'to put in' |
| | hitCaCeC | hitkarer | hitlabeʃ |
| | | 'to get a cold' | 'to dress' |
| Nouns & adjectives | | | |
| | CaCiC | karir | ʃavir |
| | | 'chilly' | 'fragile' |

| | Configuration | Reduplicated | Nonreduplicated |
|---|---|---|---|
| Nouns & adjectives | | | |
| | CaCiC | xalil<br>'flute' | sakin<br>'knife' |
| | CoCeC | ʃoded<br>'robber' | ʃomer<br>'guard' |
| | CaCCuC | ʃablul<br>'snail' | yalkut<br>'school bag' |
| | CaCCiC | ʃagrir<br>'ambassador' | ʃarvit<br>'scepter' |
| | CaCCéCet | dafdéfet<br>'paper pad' | kartéset<br>'card index' |
| | CCuCit | zxuxit<br>'glass' | kruvit<br>'cauliflower' |
| | maCCuC | maslul<br>'path' | macpun<br>'conscience' |
| | miCCaCa | mixlala<br>'college' | miʃtara<br>'police' |

Many reduplicated forms have nonreduplicated lexical counterparts (3a), but not all. There are plenty of orphan reduplicated forms, which do not have a semantically related nonreduplicated counterpart in the current stage of the language, though they may have a reduplicated counterpart in another pattern (3b).

(3) a. *Reduplicated – nonreduplicated counterparts*

| Nonreduplicated | Reduplicated | Nonreduplicated | Reduplicated |
|---|---|---|---|
| xam<br>'hot' | ximem<br>'to heat' | xamuc<br>'sour' | xamcic<br>'sour grass' |
| hed<br>'echo' | hidhed<br>'to echo' | varod<br>'pink' | vradrad<br>'pinkish' |
| daf<br>'page' | dafdéfet<br>'paper pad' | ʔavir<br>'air' | meʔavrer<br>'fan' |
| ʔec<br>'tree' | ʔacic<br>'flower pot' | kadur<br>'ball' | kidrer<br>'to dribble' |
| ʃir<br>'song' | meʃorer<br>'poet' | katan<br>'small' | katnuni<br>'petty' |
| ʔer<br>'awake' | hitʔorer<br>'to wake up' | ʃémen<br>'oil' | ʃamnuni<br>'oily' |
| lax<br>'damp' | laxluxi<br>'slightly damp' | ʃamen<br>'fat' | ʃmanman<br>'chubby' |

b. *Orphan reduplicated forms*

| | | |
|---|---|---|
| clil<br>'sound' | ʃixlel<br>'to improve' | pikpek<br>'to doubt' |
| cilcel<br>'to ring' | | |
| balal<br>'to mix' | kilel<br>'to curse' | yaʃiʃ<br>'old man' |
| bilbel<br>'to confuse' | | |
| mecamrer<br>'causing shiver' | ʔaviv<br>'spring' | parpar<br>'butterfly' |
| cmarmóret<br>'shiver' | | |
| ʔamum<br>'dim, vague' | zakik<br>'follicle' | ʃrafraf<br>'footstool' |
| ʔimʔem<br>'to dim, darken' | | |

All reduplicated forms fit into one of the four reduplication patterns, characterized by the number of different consonants (two or three) and the prosodic structure (CVCVC, CVCCVC, or CCVCCVC); the number of pairs of identical consonants (one or two) is a by-product of these two properties (see section 4).[2]

(4) *Patterns of reduplication*

| Prosodic structure | Number of pairs of identical consonants | Number of different consonants | Pattern | Examples |
|---|---|---|---|---|
| CVCVC | 1 | 2 | $C_1VC_2VC_{2C}$ | kalal<br>'to include'<br>xaviv<br>'charming' |
| CVCCVC | 1 | 3 | $C_1VC_2C_3VC_{3C}$ | sirtet<br>'to sketch'<br>ʔafsus<br>'a nobody' |

[2] Exceptions: (a) two forms with affixal reduplication (Marantz 1982), which are completely foreign to the Hebrew system: *tip-tipa* 'very little' (*tipa* 'drop') and *gev-géver* 'macho' (*géver* 'man'); (b) one form in the pattern $C_1C_2VC_3C_4VC_{4C}$, with a complex onset and only one pair of identical consonants; *smartut* 'rag'.

| Prosodic structure | Number of pairs of identical consonants | Number of different consonants | Pattern | Examples |
|---|---|---|---|---|
| CVCCVC | 2 | 2 | $C_1VC_2C_{1C}VC_{2C}$ | tiftef ‘to drip’<br>galgal ‘wheel’ |
| CCVCCVC | 2 | 3 | $C_1C_2VC_3C_{2C}VC_{3C}$ | ʃxarxar ‘blackish’<br>ʃravrav ‘plumber’ |

Reduplicated stems, like most nonreduplicated native stems, are disyllabic, conforming to the minimal and maximal bound of the minimal word constraint (see section 4 for more details on the prosodic structure).[3] Therefore, reduplication is often motivated by the requirement to form a disyllabic stem, that is, a minimal word (Bat-El 1994b, 1995, Ussishkin 1999, 2000). Indeed, monosyllabic bases often have a related reduplicated counterpart (5a). However, this cannot be the only trigger for reduplication. First, a disyllabic stem can also serve as a base for reduplication (5b), where reduplication is not forced because of a shortage in potential word structures (see the ‘‘Why not . . . ?’’ column in (5b)). Second, there is another way to expand monosyllabic stems to disyllabic ones (5c), though reduplication seems to be more common.

(5) a. *Prosodic motivation for reduplication*

| Nonreduplicated | Reduplicated | |
|---|---|---|
| kar ‘cold’ | karir ‘chilly’ | hitkarer ‘to get a cold’ |
| xag ‘holiday’ | xagiga ‘party’ | xagag ‘to celebrate’ |
| gal ‘wave’ | galil ‘cylinder’ | galgal ‘wheel’ |
| ʃen ‘tooth’ | meʃunan ‘serrated’ | ʃinanit ‘(dental) hygienist’ |
| xen ‘charm’ | xinani ‘charming’ | hitxanxen ‘to act flirtatiously’ |
| xor ‘hole’ | mexorer ‘hole puncher’ | xorer ‘to make holes’ |

[3] There are a few monosyllabic forms of the patterns $C_1C_2VC_{2C}$ and $C_1VC_{1C}$, which fit into the configurations of the few nonreduplicated monosyllabic stems. Examples are *clil* ‘sound’ and *klal* ‘rule’ (cf. *ʃvil* ‘path’ and *kfar* ‘village’);

b. *Reduplication without prosodic motivation*

| Nonreduplicated | Reduplicated | Why not . . . ? |
|---|---|---|
| ʃiger | ʃagrir | *ʃagran |
| ‘to send’ | ‘ambassador’ | (cf. ʃadxan ‘matchmaker’) |
| xamuc | xamcic | *xamac |
| ‘sour’ | ‘sour grass’ | (cf. ʔanav/ʔenav ‘grape’) |
| kadur | kidrer | *kider |
| ‘ball’ | ‘to dribble’ | (cf. kipec ‘to hop, skip’) |
| ʃakal | ʃiklel | *ʃikel |
| ‘to weigh’ | ‘to consider relative weight’ | (cf. xiʃev ‘to calculate’) |
| dover | divrer | *diver |
| ‘spokesman’ | ‘to act as a spokesman’ | (cf. yiceg ‘to represent’) |

c. *Two ways to satisfy the minimal word*[4]

| No reduplication | | Reduplication | |
|---|---|---|---|
| ʔot | ʔiyet | ʔot | ʔotet |
| ‘letter’ | ‘to spell’ | ‘sign’ | ‘to signal’ |
| nad | nayad | nad | nadad |
| ‘to move’ | ‘mobile’ | ‘to move’ | ‘to wander’ |
| ʃem | ʃiyem | ken | kinen |
| ‘name’ | ‘to name’ | ‘nest’ | ‘to nest’ |
| tag | tiyeg | dam | dimem |
| ‘label’ | ‘to label’ | ‘blood’ | ‘to bleed’ |
| tas | tayas | ʃar | meʃorer |
| ‘to fly’ | ‘pilot’ | ‘to sing’ | ‘poet’ |

I argue in section 5 that a semantic motivation is also not sufficient to predict reduplication (or the distinction between the patterns). That is, the reason for deriving a reduplicated stem can be prosodic (*kar* ‘cold’ – *karir* ‘chilly’; minimal word), semantic (*kadur* ‘ball’ – *kidrer* ‘to dribble’; repetition), or neither of these (*ʃakal* ‘to weigh’ – *ʃiklel* ‘to consider relative weight’).

*dod* ‘uncle’ and *sus* ‘horse’ (cf. *kol* ‘voice’ and *gur* ‘puppy’). The configuration of such forms, whether reduplicated or not, is not productive, though speakers identify reduplication in the $C_1C_2VC_{2C}$ pattern.

An impermissible complex onset is simplified by an epenthetic vowel, thus yielding a surface trisyllabic form. The vowel *a* appears after a historical guttural (e.g., *ʔ**a**safsuf* ‘crowd’, *x**a**macmac* ‘sourish’), and *e* after a sonorant (e.g., *l**e**vanvan* ‘whitish’).

[4] Ussishkin (1999, 2000) attributes the distinction (in denominative verbs) between a medial *y* and reduplication to the quality of the vowel; only bases with a medial *a* undergo reduplication. This analysis makes wrong predictions, given the recently coined verbs *tag* – *tiyeg* (**tigeg*) and *ken* – *kinen* (**kiyen*). The only bases that do not undergo reduplication are those with a high vowel (e.g., *kis* ‘pocket’ – *kiyes* ‘to pickpocket’).

I thus view reduplication as one of the language's strategies for stem formation.[5] The selection of a stem/word formation strategy is often lexical, and thus reduplication must be lexically specified. This is evident from the different strategies available for expressing a noun – adjective relation: apophony as in *ʃémen* 'oil' – *ʃuman* 'fat (noun)' – *ʃamen* 'fat (adjective)'; reduplication + apophony as in *tel* 'mound' – *talul* 'steep'; affixation as in *kol* 'voice' – *koli* 'vocal' and *xaver* 'friend' – *xaveri* 'friendly'; and reduplication + apophony + affixation as in *lavan* 'white' – *lavnuni* 'whitish' and *ʃémen* 'oil' – *ʃamnuni* 'oily'. These strategies for stem/word formation express the licit structural relations in the grammar, but it is often unpredictable which strategy will be selected for a given base.

With this background in mind, I now proceed to the derivation of $C_1VC_2VC_{2C}$ stems from $C_1VC_2(V)$ (e.g., *mana* 'portion' – *minen* 'to apportion'), with emphasis on the problem raised by $C_{i1}VC_{i2}VC_{j3}$ stems (e.g., *mimen* 'to finance').

## 2 The *samam* – **sasam* versus *minen* – *mimen* Problem

Greenberg's (1950) typology of the distribution of consonants in Semitic roots reflects the strong tendency of one pair of identical consonants to appear at the right edge of the stem. To use the familiar example from Arabic, forms like *samam* 'to poison' (*simem* in Hebrew) are common but those like **sasam* are very rare. McCarthy's (1979, 1981) account for this typology is based on the Obligatory Contour Principle (OCP) and the one-to-one left-to-right mapping of root consonants to prosodic units. The OCP prohibits adjacent identical root consonants, thus allowing roots like {sm}, but not like {ssm} or {smm}. One-to-one left-to-right association of root consonants with prosodic positions may leave an empty prosodic position only at the right periphery of the stem (when the number of prosodic positions exceeds the number of root consonants), allowing it to be filled by the spreading of the last consonant of the root. The surface result is a pair of identical consonants at the right edge.

In an Optimality Theory account of the *samam* – **sasam* typology, Ussishkin (2000) proposes the STRONG ANCHOR constraint, which has two requirements: (a) correspondence between the segment at $edge_i$ of the base and the segment at $edge_i$ of the output (as in the conventional ANCHOR constraint schema; McCarthy and Prince 1995, 1999), and (b) ''every'' output segment corresponding to the base segment at $edge_i$ must be located at $edge_i$ of the output. STRONG ANCHOR is violated at the left edge in $s_1am_2 \rightarrow s_{1C}as_1am_2$ and at the right edge in $s_1am_2 \rightarrow s_1am_2am_{2C}$, since a consonant at the (left/right) edge of the input has a correspondent located in the middle

[5] Note that there is no reduplication in the inflectional paradigm. This is true for contextual (syntactic) inflection, as in the inflectional paradigm of verbs, but not for inherent inflection, as in the number paradigm of nouns (see Booij 1996 and Anderson 1992 for this distinction). There are a few reduplicated broken plural nouns in Hebrew (e.g., *lev* – *levavot* 'heart(s)', *cel* – *clalim* 'shadow(s)', *cad* – *cdadim* 'side(s)'), but there is no verb that is reduplicated in one tense and nonreduplicated in another (e.g., *rac* 'he ran' – **yircac* 'he will run').

rather than at the (left/right) edge of the output ($s_1/m_2$). However, the universal ranking STRONG ANCHOR LEFT ≫ STRONG ANCHOR EDGE (where the latter refers to both edges) ensures that reduplication will appear at the right (assuming the dominance of the minimal word as a minimal and maximal bound).

(6) $s_1am_2 \rightarrow s_1am_2am_{2C}$

| $s_1am_2$ | SANCHORL | SANCHORE |
|---|---|---|
| a. ☞ $s_1am_2am_{2C}$ | | * ($m_2$) |
| b. $s_{1C}as_1am_2$ | *! ($s_1$) | * ($s_1$) |
| c. $s_1am_2as_{1C}$ | *! ($s_{1C}$) | ** ($m_2$, $s_{1C}$) |

By the Richness of the Base principle (ROTB; Prince and Smolensky 1993), any form can be an input, and it is the task of the constraint ranking to rule out illicit outputs. Indeed, the ranking in (6) would always select $s_1am_2am_{2C}$ as the optimal output, regardless of whether the input is *sam* (6), *sasam* (7a), or *samam* (7b). Since more candidates are considered below, the role of MAX (which prohibits deletion) and the OCP becomes crucial. Note that since there are no constraints on the input in Optimality Theory, the OCP must be a surface constraint that ignores an intervening vowel (but not consonant), as Rose (2000) argues for Ethiopian Semitic languages. Moreover, it refers to the base of the output; that is, the sequence $C_iVC_{iC}$, where one C is a copy of the other, does not violate the OCP, while the sequence $C_iVC_i$, where the two consonants are identical but one is not a copy of the other, does violate it (see section 3).

(7) a. $s_1as_2am_3$ (ROTB) $\rightarrow s_1am_3am_{3C}$

| $s_1as_2am_3$ | OCP | MAX | SANCHORL | SANCHORE |
|---|---|---|---|---|
| a. $s_1as_2am_3$ | *! | | | |
| b. ☞ $s_1am_3am_{3C}$ | | * ($s_2$) | | * |
| c. $s_2am_3am_{3C}$ | | * ($s_1$) | *! | * |
| d. $s_{1C}as_1am_3$ | | * ($s_2$) | *! | * |
| e. $s_1am_3as_{1C}$ | | * ($s_2$) | *! | * |

b. $s_1am_2am_3$ (ROTB) → $s_1am_{2/3}am_{2/3C}$[6]

| $s_1am_2am_3$ | OCP | MAX | SANCHORL | SANCHORE |
|---|---|---|---|---|
| a. $s_1am_2am_3$ | *! | | | |
| b. ☞ $s_1am_2am_{2C}$ | | * ($m_3$) | | * |
| c. ☞ $s_1am_3am_{3C}$ | | * ($m_2$) | | * |
| d. $s_{1C}as_1am_3$ | | * ($m_2$) | *! | * |
| e. $s_1am_3as_{1C}$ | | * ($m_2$) | *! | * |

The ranking of the OCP above MAX forces deletion of one of the identical base consonants, and the ANCHOR constraints ensure reduplication at the right edge.

Neither McCarthy nor Ussishkin takes into account stems with identical consonants at the left edge, probably because these stems are rare and thus treated as exceptions. Hebrew has a few stems of this sort (Schwarzwald 1974, Bat-El 1989), such as *mimen* 'to finance' (cf. *minen* 'to apportion'), *mimeʃ* 'to materialize' (cf. *miʃeʃ* 'to feel with the hands'), *sisgen* 'to variegate' (cf. *signen* 'to style'), *nanas* 'dwarf' (cf. *hitnoses* 'to be raised, waived'), *lulav* 'palm branch' (cf. *levavi* 'hearty'), *ʃoʃana* 'rose' (cf. *ʃinanit* '(dental) hygienist'), and *kikar* 'plaza' (cf. *karir* 'chilly').[7] As the examples in parentheses suggest, there is nothing peculiar about the segments in these stems that forces identical consonants at the left rather than at the right edge.

The list of $C_iVC_iVC_j$ stems is quite short, and it is thus tempting to leave these forms outside the analysis, that is, treat them as exceptions. Nevertheless, I claim that such forms should not be ignored, because they can freely enter the language. The source of such forms can be (a) denominative verbs derived from bases where the prefix is identical to the first consonant of the stem (e.g., *mimsar* 'relay' → *mimser* 'to transmit a signal'), (b) acronym words (e.g., *kakac* 'officer training course' from *kurs kcinim*; Bat-El 1994a, Zadok 2002), or (c) borrowings (e.g., *koktel* 'cocktail'). Note that nouns are potential bases for denominative verbs; thus, *kakac* → *kikec* 'to participate in an officer training course' and *koktel* → *kiktel* 'to prepare a cocktail' are

[6] The two optimal candidates in (7b) violate different clauses of STRONG ANCHOR. In $s_1am_2am_{2C}$ (candidate (b)), the segment at the right edge does not correspond to a segment at the right edge of the base $s_1am_2am_3$; and in $s_1am_3am_{3C}$ (candidate (c)), $m_3$ is not at the right edge, although it corresponds to the segment at the right edge of the base.

[7] Ethiopian Semitic languages also have quite a few forms with identical consonants at the left periphery (thanks to Jean-François Prunet (pers. comm.) for information and discussion). Berhane (1992) reports on 47 such verbs in Tigrinya. Leslau's (1967) work on Amharic reveals one of the sources of such forms: the pairs *qäqqärä* ~ *qäräqqärä* 'to prick the ears' and *tättägä* 'to scorch' – *tägättägä* 'to cauterize' (synchronically unrelated) suggest a historically licit reduplication followed by deletion of a medial coda. See footnote 10 for similar cases from Hebrew.

possible derivations (see Bolozky 1999). One recent coining is the verb *titel* 'to diaper', derived from a brand name for diapers, *titulim*.[8] Since the grammar of Hebrew does not reject *sasam*-type forms, the ranking in (7), where reduplication at the right edge appears regardless of the base, does not hold for Hebrew.

I claim that forms like *mimen* 'to finance' are not reduplicated. There is no form with two identical consonants at the left edge that has a related form with one occurrence of the consonant; if it has a related form, it has identical consonants in the same position (cf. *mamon* 'finance'). A language learner constructs a morphological structure on the basis of morphological relations among surface forms. Thus, in the absence of evidence for the relation between the forms $C_iVC_j$ and $C_iVC_iVC_j$, the language learner will not construct a complex structure (i.e., reduplication) for $C_iVC_iVC_j$ forms. One may argue that at the moment the learner constructs the right-edge reduplication, on the basis of the relation between $C_iVC_j$ and $C_iVC_jVC_j$, he or she may draw the generalization that identical consonants at either edge indicate reduplication. This could be a generalization encountered during acquisition, but there is evidence suggesting that it does not survive in the adult grammar.

All new words with identical consonants derived from bases with one occurrence of the consonant exhibit right-edge reduplication. Had left-edge reduplication been a nonproductive type of reduplication, we would expect at least a few new forms of this type, since nonproductive structures are also accessible. For example, the vocalic pattern {oe} is marginal and nonproductive in Hebrew, and indeed, most new verbs take the vocalic pattern {ie}. However, as shown in Bat-El 1994b, 2003a, there are a few new verbs that take the pattern {oe}, usually in free variation with {ie} (e.g., *kided* ~ *koded* 'to codify').

The experimental study reported in Berent, Shimron, and Vaknin 2001 and in Berent and Shimron 2003b further supports this claim. In this study, speakers had to produce nonce words, given a word exemplar consisting of two (and three) consonants. The most common strategy for expanding the words with two consonants was right-edge reduplication ($C_iVC_jVC_j$); left-edge reduplication ($C_iVC_iVC_j$) was marginal. These results are compatible with those obtained from nonce-probe rating experiments (Berent and Shimron 1997, 2003b), where $C_iVC_iVC_j$ forms (e.g., *liled*) were rated lower than $C_iVC_jVC_j$ (e.g., *lided*) and $C_iVC_jVC_k$ (e.g., *piled*).

The presence—and more crucially the acceptance—of (new) forms like *mimen*, which, as argued above, are nonreduplicated, suggest that the OCP is a violable constraint in Hebrew. The ranking MAX $\gg$ OCP (unlike in the ranking in (7)) allows the preservation of the two identical consonants in *mamon* $\rightarrow$ *mimen* (8a) as well as reduplication in *mana* $\rightarrow$ *minen* (8b).

[8] The word *titulim* is a child language form of *xitulim* 'diapers', whose related verb is *xitel* 'to diaper'. Thanks to Evan Cohen for providing this example.

(8) a. *mamon* 'finance' – *mimen* 'to finance'

| $m_1am_2on_3$ | MAX | OCP | SANCHORL | SANCHORE |
|---|---|---|---|---|
| a. ☞ $m_1im_2en_3$ | | * | | |
| b. $m_{1C}im_1en_3$ | *! ($m_2$) | | * | * |
| c. $m_1in_3en_{3C}$ | *! ($m_2$) | | | * |

b. *mana* 'portion' – *minen* 'to apportion'

| $m_1an_2a$ | MAX | OCP | SANCHORL | SANCHORE |
|---|---|---|---|---|
| a. ☞ $m_1in_2en_{2C}$ | | | | * |
| b. $m_{1C}im_1en_2$ | | | *! | * |
| c. $m_1in_2em_{1C}$ | | | *! | * |

Since $m_1am_2on_3$ is a possible base, by ROTB $m_1an_2on_3$ is also a possible base. As indicated by the subscripts, the identical consonants are not in a correspondence relation in either case; that is, the bases are not reduplicated. Given the ranking in (8), the output of such a base would be nonreduplicated as well, like the output of *mamon* (8a).

(9) $m_1an_2on_3$ (ROTB) – $m_1in_2en_3$ (hypothetical)

| $m_1an_2on_3$ | MAX | OCP | SANCHORL | SANCHORE |
|---|---|---|---|---|
| a. ☞ $m_1in_2en_3$ | | * | | |
| b. $m_1in_3en_{3C}$ | *! ($n_2$) | | | * |

We see, then, that forms with identical consonants at the right edge can be either reduplicated (8b) or nonreduplicated (9). We could assume that the nonreduplicated $C_iVC_jVC_j$ forms are those that do not have a $C_iVC_j$ counterpart (what I called ''orphan'' reduplicated forms). However, there are several arguments against a grammar that generates such structural ambiguity, and in favor of assigning a reduplicated structure to all $C_iVC_jVC_j$ forms, whether they have a related $C_iVC_j$ form or not.

The first argument relies on McCarthy's (2004) ''free ride'' strategy. McCarthy argues for a learning principle for morphophonemic alternations, by which the learner generalizes the mapping /A/ → [B] for all surface [B]s, whether or not they have a related /A/. A grammar with only /A/ → [B] is ''more restrictive'' than one with both /A/ → [B] and /B/ → [B] and will thus be adopted, unless the learner encounters evidence suggesting that this grammar is ''inconsistent'' with further data. According to McCarthy, consistency and greater restrictiveness are the two

requirements of grammar, which allow learners to verify whether their hypothesis should be held or rejected.

There is no evidence that such a hypothesis with regard to orphan $C_iVC_jVC_j$ forms is inconsistent with other aspects of grammar, since there are no processes of any sort that draw a distinction between orphan and nonorphan $C_iVC_jVC_j$ forms. Actually learners encounter positive evidence for this hypothesis. An orphan $C_iVC_jVC_j$ may have a related $C_iVC_jC_iVC_j$ form (e.g., **bal – balal* 'to mix' – *bilbel* 'to confuse'), just as a reduplicated $C_iVC_jVC_j$ form has (e.g., *nad* 'to move' – *nadad* 'to wander' – *nidned* 'to swing'). At the moment learners view the orphan *balal* as a reduplicated form, they can arrive at the relation between *balal* and *bilbel*, which fits into their grammar. Otherwise, if *balal* is not reduplicated and its structure is thus $C_{i1}VC_{j2}VC_{j3}$ (where the identical consonants are not in correspondence), the structure of the related form *bilbel* would be $C_{i1}VC_{j2}C_{i1C}VC_{j3}$. Not only do learners not have surface relations to support this structure (as in a hypothetical *kalam – kilkem*), the additional reduplicated structure, which unlike the others is not limited to the right edge, reduces the restrictiveness of their grammar. I thus argue that in the presence of positive evidence for reduplicated structure provided by reduplicated forms with an existing base, learners construct a grammar that assigns a reduplicated structure to all forms with identical consonants at the right edge.[9]

Further support for the ''free ride'' taken by orphan reduplicated forms can be drawn from experimental studies. Of particular interest are the experiments involving lexical decision tasks reported by Berent, Shimron, and Vaknin (2001), where participants had to determine whether a given word is an actual word in Hebrew. The list consisted of actual and nonce words of three types: words with two identical consonants at the left edge ($C_iVC_iVC_j$), two identical consonants at the right edge ($C_iVC_jVC_j$), and no identical consonants ($C_iVC_jVC_k$). There were significant differences in response latency with respect to the three types of nonce words (but none with respect to the actual words). $C_iVC_iVC_j$ forms were rejected significantly faster than $C_iVC_jVC_j$ forms, which means that they were immediately identified as nonwords. More significant for the argument regarding the reduplicated structure of orphan forms is that $C_iVC_jVC_j$ forms were rejected significantly more slowly than $C_iVC_jVC_k$ forms. As Berent, Shimron, and Vaknin emphasize, frequency does not play a role here since the less frequent forms ($C_iVC_jVC_j$) were rejected more slowly. This result suggests that forms with identical consonants at the right edge ($C_iVC_jVC_j$) are perceived as structurally complex, that is, reduplicated. Novel morphological complex words take longer to discriminate than noncomplex words since discrimination involves structural decomposition.

Having argued that Hebrew speakers can parse reduplicated forms without reference to a base, I now propose the parsing constraints that allow speakers to distinguish between reduplicated and nonreduplicated stems with identical consonants, regardless of their status in the lexicon.

[9] Recall that I focus here on parsing surface forms and thus make no claims about whether the learner establishes an abstract underlying base for the orphan reduplicated forms.

## 3 Parsing Stems with Identical Consonants

In this section, I present the linguistic information required to parse stems with identical consonants. A stem with *n* different consonants has $n^2$ logically possible forms with one pair of identical consonants (I will discuss two pairs of identical consonants later on). Since the position of the identical consonants is relevant, rather than their content, $C_iVC_jVC_j$ also stands for $C_jVC_iVC_i$, $C_hVC_iC_jVC_h$ for $C_jVC_iC_hVC_j$, and so on.

(10) a. *A stem with two different consonants*

i. **$C_iVC_jVC_j$**
ii. $C_iVC_iVC_j$
iii. $C_iVC_jVC_i$
iv. $C_jVC_iVC_j$

b. *A stem with three different consonants*

i. **$C_hVC_iC_jVC_j$**
ii. $C_hVC_jC_iVC_j$
iii. $C_jVC_hC_iVC_j$
iv. $C_hVC_iC_jVC_h$
v. $C_hVC_iC_hVC_j$
vi. $C_hVC_hC_iVC_j$
vii. $C_hVC_iC_jVC_i$
viii. $C_hVC_iC_iVC_j$
ix. $C_iVC_hC_iVC_j$

For both (10a) and (10b), only the first form represents a reduplicated stem. The other forms also exist in Hebrew, but, as argued above, they are not reduplicated. That is, while there are stems like *mimen* 'to finance' (10aii), *ʃóreʃ* 'root' (10aiii), *diskes* 'to discuss' (10bii), and *sibsed* 'to subsidize' (10bv), such stems do not have a semantically related counterpart like **man*, **ʃar*, **dasak*, and **sabad*, respectively.[10]

The constraints allowing speakers to distinguish between reduplicated and nonreduplicated stems with identical consonants refer to edges of two domains, a surface base and a stem. In a reduplicated stem (11a), the base is nested within the stem, and the copy resides outside the base. In a nonreduplicated stem (11b), the edges of the base and the stem coincide (''{ }'' marks the edges of the base and ''[ ]'' the edges of the stem).

(11) a. The domain structure of a reduplicated stem: $[\{ \ldots \}_{\text{Base}} \ldots ]_{\text{Stem}}$
b. The domain structure of a nonreduplicated stem: $[\{ \ldots \}_{\text{Base}}]_{\text{Stem}}$

[10] Actually, some speakers identify a relation between *ʃʃofére<u>t</u>* 'tube' and *ʃofar* 'ceremonial ram's horn'. The interesting point is that many speakers say *ʃʃorfére<u>t</u>* (i.e., with an additional *r*), which is a licit reduplicated form (also *metultél<u>et</u>*/*metutél<u>et</u>* 'pendulum', *ʃolʃél<u>et</u>*/*ʃoʃél<u>et</u>* 'dynasty'; but *xacocra*/**xacorcra* 'trumpet', where reduplication would result in a medial complex onset in the last syllable). The semantic relation and the structural adjustment are both historically valid; *ʃofar* > **ʃʃurpére<u>t</u>* (reduplication) > *ʃʃofére<u>t</u>* (deletion of a liquid in a nonfinal coda position).

As shown in (2), reduplicated and nonreduplicated stems take the same configurations (prosodic structures, vocalic patterns, and affixes). Thus, the notion ''stem'' refers here to a word without its prefixes and suffixes; crucially, the copy is part of the stem rather than a suffix. The base is not necessarily an existing (abstract or surface) form, but it is certainly a unit that speakers can arrive at, given the constraints to be proposed below. We thus have to assume a principle defining a base.

(12) *Basehood*
A base consists of all and only base segments.

Since the position of the vowels is determined by higher-ranked constraints on syllabic structure, Basehood has an effect only on the consonants. Basehood selects [{$C_1VC_2$}$VC_{2C}$] as a licit structure, but not *[{$C_1VC_{2C}$}$VC_2$], where a base segment ($C_2$) is outside the base, or *[{$C_1VC_2VC_{2C}$}], where a nonbase segment ($C_{2C}$) is within the base.

In addition to Basehood (which will not be considered in the tableaux below), two constraints are required for a correct parsing of stems with identical consonants: Surface Correspondence by Identity and Surface Correspondence by Position. These two constraints define the conditions under which two consonants are in a correspondence relation. I use the term *correspondence* here to refer to a pair of identical consonants one of which is the copy of the other.

(13) *Surface Correspondence by Identity* (*SCorrI*)
If S is a stem,
$C_x$ & $C_y \in$ S, and
$C_x$ & $C_y$ are identical,
Then $C_x$ & $C_y$ are correspondents.

This constraint prohibits identical consonants in a stem unless they are correspondents. That is, if the stem has identical consonants, it is reduplicated, one consonant being the copy of the other.

Rose and Walker (2001) argue that all segments in a word are in a correspondence relation (not in the morphological sense I am using here) and that correspondence among similar consonants activates long-distance agreement (i.e., assimilation/harmony). However, while some languages exhibit long-distance agreement, which may result in identical consonants, others do not allow identical consonants in a word. To account for this diversity, Yip (1998) proposes two competing constraints: a revised version of the OCP, stating that the ''[o]utput must not contain two identical elements,'' and Repeat, stating that the ''[o]utput must contain two identical elements'' (p. 221). In both studies, adjacency plays a role, such that the closer the corresponding identical/similar segments, the more likely they are to be affected (see Pierrehumbert 1993 and Frisch, Pierrehumbert, and Broe 2004 for the relation between distance and similarity/identity, and Berent and Shimron 2003a for crucial distinctions between similarity and identity). The diversity among languages in this respect may fit into a scale: similarity $\Rightarrow$ phonological identity $\Rightarrow$ morphological identity (where morphological identity means reduplication). Each of these structures is rejected by some constraint, and when the constraint is sufficiently high in the relevant

ranking, the structure changes (not necessarily synchronically) following the scale (though the option of dissimilation is also available).

Here, I look at phonological identity, asking under which conditions it is parsed as morphological identity. SCORRI states that identical (but not similar) consonants are in a correspondence relation.[11] This activates the assignment of a domain structure (11), placing the identical consonants in different domains and thus interpreting the phonological identity as morphological identity. However, the assignment of a domain is restricted by SURFACE CORRESPONDENCE BY POSITION.

(14) *SURFACE CORRESPONDENCE BY POSITION (SCORRP)*

If S is a stem,
$C_x$ & $C_y \in$ S,
$C_x$ & $C_y$ are identical, and
$C_x$ & $C_y$ are at the right edges of the domains,
Then $C_x$ & $C_y$ are correspondents.

While SCORRI assigns correspondence to every pair of identical consonants, regardless of their position in the stem, SCORRP restricts corresponding segments to the right edge. SCORRP is more restrictive than SCORRI and therefore can have an effect only if it is ranked higher in the hierarchy.[12] Given the ranking SCORRP $\gg$ SCORRI, in positions other than the right edge, identity does not imply morphological correspondence, and the forms are thus not reduplicated. As noted by a reviewer, SCORRI seems to be unnecessary, as it is included in SCORRP. However, identity has an independent status in the grammars of languages, and the positional restriction must be viewed as an additional restriction rather than as an integral part of identity.

As shown in the tableaux in (15), these two constraints allow speakers to assign the correct structure to forms with identical consonants and thus to arrive at a base without reference to some base in the lexicon (reference to the lexicon is, however, required for semantic purposes). Therefore, every stem with identical consonants at the right edge will be parsed as reduplicated, whether it is an orphan reduplicated form, a nonce word, or a reduplicated form that has a nonreduplicated counterpart; stems with identical consonants in other positions will be parsed as nonreduplicated. In the parsing tableaux below, the input is a surface form heard by the listener and the output represents the structure the listener assigns to this form.

[11] I ignore the few reduplicated forms where the corresponding consonants differ in the value of [continuant] (e.g., *sibev* 'to turn', *xibev* 'to like', *hafaxpax* 'fickle'). The stop-fricative alternation is unstable in the current stage of Hebrew (see Adam 2002), and the difference in the value of [continuant] is often eliminated in the colloquial register, in favor of the fricative (thus, *sivev*, *xivev*, *hafaxfax*).

[12] Everett and Berent (1997), who propose *INITIAL IDENTITY $\gg$ *IDENTITY, and Rose (2000) view stems with identical consonants at the right edge (i.e., reduplicated forms) as violating the OCP/IDENTITY. However, these studies are concerned with phonological structure, rather than morphological parsing, and therefore do not consider the distinction between reduplicated and nonreduplicated candidates with two identical consonants at the right edge (e.g., [{$k_1 i d_2$}$ed_{2C}$] vs. *[{$k_1 i d_2 e d_3$}]; see candidates (a) and (c) in (15ai)). This distinction is crucial, given the presence of [$m_1 i${$m_2 e n_3$}] but the absence of *[{$m_{1C} i m_1 e n_2$}].

(15) *Parsing stems with one pair of identical consonants*[13]

a. A pair of identical consonants at the right periphery

i. $C_iVC_jVC_j \Rightarrow C_1VC_2VC_{2C}$ (*kided* 'to codify')

| kided | SCORRP | SCORRI |
|---|---|---|
| a. ☞ [{$k_1id_2$}$ed_{2C}$] | | |
| b. [{$k_1id_{2C}ed_2$}] | *! | |
| c. [{$k_1id_2ed_3$}] | | *! |

ii. $C_hVC_iC_jVC_j \Rightarrow C_1VC_2C_3VC_{3C}$ (*cixkek* 'to giggle')

| cixkek | SCORRP | SCORRI |
|---|---|---|
| a. ☞ [{$c_1ix_2k_3$}$ek_{3C}$] | | |
| b. [{$c_1ix_2k_{3C}ek_3$}] | *! | |
| c. [{$c_1ix_2k_3ek_4$}] | | *! |

b. A pair of identical consonants at the left periphery

i. $C_iVC_iVC_j \Rightarrow C_1VC_2VC_3$ (*mimen* 'to finance')

| mimen | SCORRP | SCORRI |
|---|---|---|
| a. [$m_{1C}i${$m_1en_2$}] | *! | |
| b. [{$m_1im_{1C}en_2$}] | *! | |
| c. ☞ [{$m_1im_2en_3$}] | | * |

ii. $C_hVC_hC_iVC_j \Rightarrow C_1VC_2C_3VC_4$ (*sisgen* 'to variegate')

| sisgen | SCORRP | SCORRI |
|---|---|---|
| a. [$s_{1C}i${$s_1g_2en_3$}] | *! | |
| b. [{$s_1is_{1C}g_2en_3$}] | *! | |
| c. ☞ [{$s_1is_2g_3en_4$}] | | * |

[13] Candidate (b) in all the tableaux, except (15ci), is also ruled out by BASEHOOD (12).

c. A pair of identical consonants interrupted by another consonant

i. $C_iVC_jVC_i \Rightarrow C_1VC_2VC_3$ (*ʃileʃ* 'to triple')

| ʃileʃ | SCorrP | SCorrI |
|---|---|---|
| a. $[\int_{2C}i\{l_1e\int_2\}]$ | *! | |
| b. $[\{\int_1il_2\}e\int_{1C}]$ | *! | |
| c. ☞ $[\{\int_1il_2e\int_3\}]$ | | * |

ii. $C_hVC_iC_hVC_j \Rightarrow C_1VC_2C_3VC_4$ (*safsal* 'bench')

| safsal | SCorrP | SCorrI |
|---|---|---|
| a. $[s_{2C}a\{f_1s_2al_3\}]$ | *! | |
| b. $[\{s_1af_2s_{1C}al_3\}]$ | *! | |
| c. ☞ $[\{s_1af_2s_3al_4\}]$ | | * |

iii. $C_hVC_iC_jVC_i \Rightarrow C_1VC_2C_3VC_4$ (*xicrec* 'to play the trumpet')

| xicrec | SCorrP | SCorrI |
|---|---|---|
| a. $[\{x_1ic_2r_3\}ec_{2C}]$ | *! | |
| b. $[\{x_1ic_{3C}r_2ec_3\}]$ | *! | |
| c. ☞ $[\{x_1ic_2r_3ec_4\}]$ | | * |

SCorrI is violated when $C_i$-$C_i$ are not $C_i$-$C_{iC}$ (or $C_{iC}$-$C_i$), which means that identical consonants in the stem are not correspondents (i.e., one is not the copy of the other, as indicated by the absence of the subscript *c*). Given the corresponding segments $C_i$-$C_{iC}$, SCorrP is violated when $C_{iC}$ (the copy), $C_i$ (the copied segment), or both are not at the right edge of a domain.

The constraint ranking proposed above refers to the surface forms, but it also has an effect on the structure of the words in the lexicon. Given this constraint ranking, every form with identical consonants at the right edge will be parsed as reduplicated (see (15a–b)). Lexicon Optimization (Prince and Smolensky 1993, Inkelas 1994, Itô, Mester, and Padgett 1995), by which the learner selects an input on the basis of the output and the constraint ranking, will then require selecting a reduplicated form as an input. That is, although by ROTB a nonreduplicated $C_{i1}VC_{j2}VC_{j3}$ is a possible input, Lexicon Optimization eliminates it as an input in the language, given the constraint ranking proposed above.

As reviewed in section 1, Hebrew also has reduplicated forms with two pairs of identical consonants. A few examples are given in (16).

(16) *Reduplicated forms with two pairs of identical consonants*

a. Bases with two different consonants: $C_1VC_2C_{1C}VC_{2C}$

| Nonreduplicated base | Reduplicated | Orphan reduplicated |
|---|---|---|
| kav | kivkev | sixsex |
| ‘line’ | ‘to mark with line’ | ‘to cause disagreement’ |
| yafe | yafyuf | barbur |
| ‘beautiful’ | ‘beautiful (derogatory)’ | ‘swan’ |

b. Bases with three different consonants: $C_1C_2VC_3C_{2C}VC_{3C}$

| Nonreduplicated base | Reduplicated | Orphan reduplicated |
|---|---|---|
| kélev | klavlav | ʃravrav |
| ‘dog’ | ‘little dog’ | ‘plumber’ |
| ktana | ktantónet | cmarmóret |
| ‘small (fem.)’ | ‘very small (fem.)’ | ‘shiver’ |

The constraint ranking proposed above does not account for reduplicated forms with two pairs of identical consonants, since only the two members of one pair can occupy the right edges of the domains, and therefore the other pair violates SCorrP (‘‘*☞’’ indicates the wrong optimal candidate, and ‘‘√’’ the actual form).

(17) $C_iVC_jC_iVC_j \Rightarrow {}^*C_1VC_2C_3VC_4$ (*kivkev* ‘to mark with line’)

| kivkev | SCorrP | SCorrI |
|---|---|---|
| a. √ $[\{k_1iv_2\}k_{1C}ev_{2C}]$ | *! $(k_1\text{-}k_{1C})$ | |
| b. $[k_{1C}iv_{2C}\{k_1ev_2\}]$ | *! $(k_{1C}\text{-}k_1, v_{2C})$ | |
| c. $[\{k_1iv_2k_3\}ev_{2C}]$ | *! $(v_2)$ | * |
| d. *☞ $[\{k_1iv_2k_3ev_4\}]$ | | *!* |

It is thus necessary to expand SCorrP, so that it will refer to two pairs of identical consonants, rather than one. Notice that in reduplicated forms with one pair of identical consonants, only a vowel can intervene between the two identical consonants. In reduplicated forms with two pairs of identical consonants, the intervening material consists of a vowel and a consonant. However, the intervening consonant is always a member of the other pair of identical consonants. Thus, as

modified in (18), the first pair is evaluated with respect to the right edges of the domains, as in the earlier version, and the second with respect to the first one (the addition to the version in (14) is italicized).

(18) *Surface Correspondence by Position (SCorrP)*

If S is a stem,
*$C_a$ & $C_b$ and* $C_x$ & $C_y$ ∈ S,
*$C_a$ & $C_b$ are identical and* $C_x$ & $C_y$ are identical, and
*$C_a$ & $C_b$ are correspondents and* $C_x$ & $C_y$ are correspondents,
Then $C_x$ & $C_y$ are at the right edges of the domains, and
*X in $C_bXC_y$ and $C_aXC_x$ does not include a consonant.*[14]

Given the expanded version in (18), candidate (a) (like candidate (d)) is not ruled out by SCorrP in (19a–b), and SCorrI gets to select the reduplicated candidate (candidate (a)) as the optimal one. When a consonant not belonging to a pair of identical consonants intervenes between the identical consonants, as in (19c), SCorrP rules out the reduplicated structure (candidate (a)), and the nonreduplicated one (candidate (b)) is then the optimal one.

(19) a. $C_iVC_jC_iVC_j \Rightarrow C_1VC_2C_{1C}VC_{2C}$ (*kivkev* 'to mark with line')

| kivkev | SCorrP | SCorrI |
|---|---|---|
| a. ☞ $[\{k_1iv_2\}k_{1C}ev_{2C}]$ | | |
| b. $[k_{1C}iv_{2C}\{k_1ev_2\}]$ | *! | |
| c. $[\{k_1iv_2k_3\}ev_{2C}]$ | *! | * |
| d. $[\{k_1iv_2k_3ev_4\}]$ | | *!* |

b. $C_hC_iVC_jC_iVC_j \Rightarrow C_1C_2VC_3C_{2C}VC_{3C}$ (*klavlav* 'little dog')

| klavlav | SCorrP | SCorrI |
|---|---|---|
| a. ☞ $[\{k_1l_2av_3\}l_{2C}av_{3C}]$ | | |
| b. $[\{k_1l_{2C}av_{3C}l_2av_3\}]$ | *! | |
| c. $[\{k_1l_2av_3l_4a\}v_{3C}]$ | *! | * |
| d. $[\{k_1l_2av_3l_4av_5\}]$ | | *!* |

[14] The fact that X is always a vowel is due to restrictions on prosodic structure. A form like **kilvlev*, which is ill formed because of its prosodic structure (see section 4), also respects SCorrP.

c. $C_hC_iVC_jC_kVC_j \Rightarrow C_1C_2VC_3C_4VC_5$ (*klavtav* 'a court reporter', a blend of *klavlav* 'little dog' and *katav* 'reporter')

| klavtav | SCORRP | SCORRI |
|---|---|---|
| a. [{$k_1l_2av_3$}$t_4av_{3C}$] | *! | |
| b. ☞ [{$k_1l_2av_3t_4av_5$}] | | * |

So far, I have answered the first question addressed in the introduction: how do speakers distinguish between reduplicated and nonreduplicated stems with identical consonants, without reference to a base? Now I turn to the second question: what is the system responsible for the patterns of reduplication?

## 4 The Patterns of Reduplication

There are four patterns of reduplication (see (4)): $C_1VC_2VC_{2C}$, $C_1VC_2C_3VC_{3C}$, $C_1VC_2C_{1C}VC_{2C}$, and $C_1C_2VC_3C_{2C}VC_{3C}$. Their distinctive properties are their prosodic structure (CV, CVC, or CCVC as the first syllable) and their number of different consonants (two or three). Arguments supporting a unified account for all the patterns are given in section 1 (prosodic) and section 5 (semantic). Here, I show that all these patterns can be derived by the same system.

I assume that all types of reduplication are triggered by a morphological constraint COPY, which, like REPEAT in Yip 1998, requires every segment to appear twice in the stem (see section 6 for further discussion). COPY is fully respected in total reduplication, and it gets a violation mark for every segment that appears only once in the stem, as in the base-reduplicant faithfulness relation proposed by McCarthy and Prince (1995, 1999). However, unlike MAX-BR, which requires base-reduplicant faithfulness, COPY (like REPEAT) does not make reference to a base. The effect of COPY on vowels cannot be surface true, since the surface vowels have to be selected from a closed set of vocalic patterns (see Bat-El 2003a,b). I thus ignore the vowels in the tableaux below.

Most surface forms are disyllabic (see footnote 3 for exceptions), thus respecting the set of constraints defining the minimal word as the minimal and maximal bound (see Ussishkin 2000 and references therein). In addition, the following ranking of prosodic constraints is relevant:

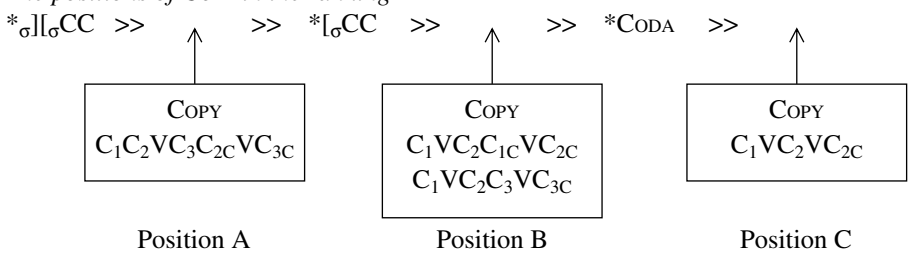
(20) $*_\sigma][_\sigma CC \gg *[_\sigma CC \gg *$CODA
  a. $*_\sigma][_\sigma CC$: A medial syllable does not have a complex onset.
  b. $*[_\sigma CC$: An initial syllable does not have a complex onset.
  c. *CODA: A syllable does not have a coda.

The effect of these constraints in Hebrew is evident beyond reduplication. Evidence for $*_\sigma][_\sigma CC$ can be drawn from the verb paradigm: a vowel in a stem-final syllable is deleted when a vowel-initial suffix is added (e.g., *gidel – gidla* 'he – she raised'), unless the resulting form would have a medial complex onset (e.g., *tilfen – tilfena*, **tilfna* 'he – she phoned'). The ranking

$*_{\sigma}][_{\sigma}CC \gg *[_{\sigma}CC$ is motivated by the fact that $*_{\sigma}][_{\sigma}CC$ blocks deletion in verbs to avoid a medial complex onset, but $*[_{\sigma}CC$ does not block deletion in adjectives, where the stem vowel in the first (rather than the last) syllable is deleted when a suffix is added (e.g., *ʃavir – ʃvirim* 'fragile sg. – pl.'). Evidence for $*[_{\sigma}CC$ can be drawn from the fact that verbs, unlike nouns and adjectives, do not allow complex onsets anywhere in the stem (and this is the reason why there are no verbs in the reduplicated pattern $C_1C_2VC_3C_{2C}VC_{3C}$).[15] Therefore, verbs with four consonants are syllabified CVCCVC, rather than CCVCVC. *CODA is violated by the first syllable in CVCCVC and $*[_{\sigma}CC$ is violated by the first syllable in CCVCVC; therefore, the ranking is $*[_{\sigma}CC \gg$ *CODA. Despite being massively violated (because of its relatively low position in the hierarchy), *CODA shows its effect in disyllabic stems with three consonants, which are usually syllabified CVCVC rather than CVCCV. Both forms violate *CODA, but the former is preferred owing to the constraint FINALC, which requires stems to end in a consonant (McCarthy 1993). As shown in Bat-El 1994a, these constraint rankings have a strong effect on acronyms, which never allow complex onsets, either in initial or in medial position, and which prefer a final coda over a medial one.

To account for the different prosodic structures of the reduplicated stems, I propose that the constraint COPY has three possible positions in the ranking given in (20). This proposal is in the spirit of studies on variation, among dialects (Yip 1996), sublexicons (Itô and Mester 1995), and cophonologies (Anttila 2002).

(21) *The positions of COPY in the ranking*




This model provides three different prosodic structures (given the minimal word as a minimal and maximal bound): CCVCCVC, CVCCVC, and CVCVC. The existence of four patterns of reduplication is due to the number of different consonants. I assume that COPY is not iterative, and thus every consonant can have only one copy. Thus, a reduplicated CVCVC stem can host only two different consonants ($C_1VC_2VC_{2C}$), a reduplicated CVCCVC stem can host two or three different consonants ($C_1VC_2C_{1C}VC_{2C}$, $C_1VC_2C_3VC_{3C}$), and a reduplicated CCVCCVC stem can host three different consonants ($C_1C_2VC_3C_{2C}VC_{3C}$). The tableaux in (22) demonstrate this inventory of patterns.

[15] This restriction does not hold for denominative verbs, which have to preserve all the base consonants, thus allowing complex onsets anywhere in the stem; for example, *télegraf* 'telegraph' → *tilgref* 'to telegraph', *traklin* 'salon' → *triklen* 'to tidy up', *ʔabstrákti* 'abstract' → *ʔibstrekt* 'to make something abstract' (Bolozky 1978b, McCarthy 1984, Bat-El 1994b, 1995). Thus, a verb derived from the nominal/adjectival pattern $C_1C_2VC_3C_{2C}VC_{3C}$ preserves the complex onset (e.g., *sxarxóret* 'dizziness' → *sxirxer* 'to cause dizziness'; but *sixrer* is more common).

(22) a. *Three different consonants*

i. $C_1C_2VC_3C_{2C}$ $VC_{3C}$: $*_\sigma][_\sigma CC$ >> **COPY** >> $*[_\sigma CC$

| | $*_\sigma][_\sigma CC$ | **COPY** | $*[_\sigma CC$ | *CODA |
|---|---|---|---|---|
| a. vrad.vrad | *! | | ** | ** |
| b. ☞ vrad.rad | | * | * | ** |
| c. var.drad | *! | * | * | ** |
| d. var.dad | | **! | | ** |
| e. va.rad | | **!* | | * |

ii. $C_1VC_2C_3VC_{3C}$: $*[_\sigma CC$ >> **COPY** >> *CODA

| | $*_\sigma][_\sigma CC$ | $*[_\sigma CC$ | **COPY** | *CODA |
|---|---|---|---|---|
| a. dvir.dver | *! | ** | | ** |
| b. dvir.ver | | *! | * | ** |
| c. div.rver | *! | * | * | ** |
| d. ☞ div.rer | | | ** | ** |
| e. di.ver | | | ***! | * |

b. *Two different consonants*

i. $C_1VC_2C_{1C}VC_{2C}$: **COPY** >> *CODA ($*[_\sigma CC$ and COPY are not crucially ranked)

| | $*_\sigma][_\sigma CC$ | $*[_\sigma CC$ | **COPY** | *CODA |
|---|---|---|---|---|
| a. kvi.kev | | *! | | * |
| b. ☞ kiv.kev | | | | ** |
| c. ki.vev | | | *! | ** |

ii. $C_1VC_2VC_{2C}$: *CODA >> **COPY**

| | $*_\sigma][_\sigma CC$ | $*[_\sigma CC$ | *CODA | **COPY** |
|---|---|---|---|---|
| a. kdi.ked | | *! | * | |
| b. kid.ked | | | **! | |
| c. ☞ ki.ded | | | * | * |

As shown above, the selection of a particular pattern depends on the number of different consonants. This, however, leaves two options for each case.

(23) a. 3 different consonants: positions A and B in (21) for nouns, but position B only for verbs (see above for syllabic restrictions on verbs)
  i. $C_1C_2VC_3C_{2C}VC_{3C}$ levanvan A 'whitish (usually inanimate)' (*e* epenthetic)
  ii. $C_1VC_2C_3VC_{3C}$ lavnuni B 'whitish (usually animate)'
b. 2 different consonants: positions B and C in (21)
  i. $C_1VC_2C_{1C}VC_{2C}$ nidned B 'to swing'
  ii. $C_1VC_2VC_{2C}$ nadad C 'to wander'

In the following section, I consider the semantic properties that have been suggested for distinguishing between the two patterns in each case. I show that in light of a wide body of data, like that considered here, the semantic properties do not have predictive power, though they may lend themselves to an a posteriori partial generalization. The absence of a semantic distinction supports a unified account of the different patterns of reduplication, as proposed above. It also requires stems to be specified not only for COPY, but also for its position in the ranking. As a result, this model expresses the licit patterns of reduplication in Hebrew and the limits on deriving a new reduplicated form.

## 5 The Role of Semantics in Hebrew Reduplication

Studies of Hebrew reduplication have suggested that some patterns of reduplication are associated with semantic properties. McCarthy (1979, 1981) notes that Tiberian Hebrew's verbal pattern $C_1C_2VC_3C_{2C}VC_{3C}$ designates intensification; in Modern Hebrew, this pattern, which exists only in nouns and adjectives, is claimed to designate diminutive (Bolozky 1999, Graf 2002). Ussishkin (1999, 2000) supports his formal distinction between the verbal patterns $C_1VC_2C_{1C}VC_{2C}$ (RED suffixation) and $C_1VC_2VC_{2C}$ (STRONG ANCHOR) by claiming that the former designates durative or repetitive meaning while the latter is semantically neutral.

When a larger and more diverse body of data is considered, it appears that the semantic distinction between the patterns does not always hold. As the examples in (24) show, stems in all patterns may be semantically neutral or may designate one of the properties mentioned above (i.e., diminutive for nouns/adjectives and repetitive/durative for verbs).

(24) a. *Nouns and adjectives*

| | Diminutive | | Neutral | |
|---|---|---|---|---|
| Pattern | Reduplicated | Related form | Reduplicated | Related form |
| $C_1VC_2VC_{2C}$ | dagig 'little fish' | dag 'fish' | srara 'authority' | sar 'minister' |
| $C_1VC_2C_{1C}VC_{2C}$ | salsala 'little basket' | sal 'basket' | galgal 'wheel' | gal 'wave' |
| $C_1VC_2C_3VC_{3C}$ | kamtut 'small wrinkle' | kémet 'wrinkle' | saharuri 'lunatic' | sáhar 'moon' |
| $C_1C_2VC_3C_{2C}VC_{3C}$ | klavlav 'little dog' | kélev 'dog' | ʔasafsuf 'crowd' | ʔósef 'collection' |

b. *Verbs*

| | Durative/Repetitive | | Neutral | |
|---|---|---|---|---|
| Pattern | Reduplicated | Related form | Reduplicated | Related form |
| $C_1VC_2VC_{2C}$ | dimem 'to bleed' | dam 'blood' | cided 'to side with' | cad 'side' |
| $C_1VC_2C_{1C}VC_{2C}$ | cilcel 'to ring' | clil 'sound' | bilbel[16] 'to confuse' | blil 'mixture' |
| $C_1VC_2C_3VC_{3C}$ | kidrer 'to dribble' | kadur 'ball' | divrer 'to act as a spokesman' | dover 'spokesman' |

I claim that structural similarity between semantically related forms does not necessarily indicate that it is the structure that carries the shared semantic property. A similar state of affairs is found with nouns ending in the nominal suffix *-on*, which appears in derived nouns with a wide variety of semantic properties.[17] While *-on* does not assign a specific semantic property (25a), it consistently appears in nouns denoting some type of newspaper (25b).

(25) a. *The different semantic properties associated with nouns ending in* -on

| Related word | *-on* noun | Related word | *-on* noun |
|---|---|---|---|
| ʃaʔa 'hour' | ʃaʔon 'clock' | ʔavir 'air' | ʔaviron 'airplane' |
| ʔiver 'blind' | ʔivaron 'blindness' | ʃabat 'Sabbath' | ʃabaton 'sabbatical' |
| mila 'word' | milon 'dictionary' | halixa 'walking' | halixon 'treadmill' |
| mazon 'food' | miznon 'buffet' | ʃalat 'to rule' | ʃilton 'government' |

b. *Nouns with* -on *denoting some type of newspaper (np.)*

| Related word | *-on* noun | Related word | *-on* noun |
|---|---|---|---|
| ʔet 'time' | ʔiton 'newspaper' | ʃavúa 'week' | ʃvuʔon 'weekly np.' |
| makom 'place' | mekomon 'local np.' | yérax 'month' | yarxon 'monthly np.' |
| ʔale 'leaf' | ʔalon 'leaflet' | bite 'to pronounce' | bitaʔon 'ideological np.' |

[16] Notice that contrary to Ussishkin's (1999, 2000) prediction, *balal* 'to mix', rather than *bilbel* 'to confuse', has a durative meaning.

[17] This *-on* is not the diminutive *-on* found in *gamad* 'dwarf' – *gamadon* 'little dwarf' (see Bat-El 1997).

Despite the examples in (25b), I am reluctant to conclude that one of the many semantic properties of *-on* is 'disposable reading material'. Rather, I claim that *-on* is a general nominal suffix, with no specific semantic property. The semantic property shared by the forms in (25b) is based on the generic word *ʔiton* 'newspaper', rather than on the suffix *-on*. Occasionally, a word takes the structure of another word in order to reflect some semantic affiliation that is not expressed by a shared base or by a semantically specified structure. The noun *mefagéa* 'suicide bomber', for example, takes the configuration of the older word *mexabel* 'terrorist'. Another possible form for 'suicide bomber' could be **pagʔan*, but it has not been chosen since it is structurally similar to *xablan* 'bomb disposal expert', which has a positive connotation. Similarly, *kapliya* 'caplet' – *tavliya* 'tablet', *yevu* 'import' – *yecu* 'export', *gizʔanut* 'racism' (cf. *géza* 'race', *gizʔan* 'racist') – *minanut* 'sexism' (cf. *min* 'sex', **minan*), *toxna* 'software' – *gonva* 'pirated software' (cf. *ganav* 'to steal'), and the name of a taxi company *ʃaronit* (where *ʃaron* is the name of the area where it operates), based on *monit* 'taxi'. Of course, the common structure of the forms in each pair does not carry the shared semantic property, as there are plenty of forms using these structures that do not carry this meaning. Similarly, the common pattern of reduplication of some diminutive nouns/adjectives or some repetitive/durative verbs does not carry the semantic property.

Further justification for this argument is required for the pattern $C_1C_2VC_3C_{2C}VC_{3C}$, because unlike the other patterns, it appears in forms the majority of which indeed share a meaning: diminutive. However, the predictive power of this semantic property is weak, since bases with three consonants can select $C_1VC_2C_3VC_{3C}$ as well, sometimes as an alternative to $C_1C_2VC_3C_{2C}VC_{3C}$.

(26) *Diminutive reduplicated forms with three different consonants*

| Base | $C_1VC_2C_3VC_{3C}$(i) | $C_1C_2VC_3C_{2C}VC_{3C}$ |
|---|---|---|
| Nouns | | |
| ʃéver 'breakage' | ʃavrir 'splint' | |
| xélek 'part' | xelkik 'particle' | |
| kémet 'wrinkle' | kamtut 'little wrinkle' | |
| géver 'man' | | gvarvar 'macho' |
| zakan 'beard' | | zkankan 'little beard' |
| Adjectives | | |
| ʔagum 'gloomy' | ʔagmumi 'gloomy' | |
| yaʃen 'sleeping' | yaʃnuni 'sleepy' | |

| Base | $C_1VC_2C_3VC_{3C}$(i) | $C_1C_2VC_3C_{2C}VC_{3C}$ |
|---|---|---|
| ʃamen | | ʃmanman |
| 'fat' | | 'chubby' |
| xalak | | xalaklak |
| 'smooth' | | 'smooth' |
| varod | | vradrad |
| 'pink' | | 'pinkish' |
| ʃaxor | | ʃxarxar |
| 'black' | | 'blackish' |
| katan | katnuni | ktantan |
| 'small' | 'petty' | 'very small' |
| ʔadom | ʔadmumi | ʔadamdam |
| 'red' | 'reddish' | 'reddish' |
| lavan | lavnuni | levanvan |
| 'white' | 'whitish' | 'whitish' |

There are also structural properties that may support the claim that $C_1C_2VC_3C_{2C}VC_{3C}$ is different from the other patterns of reduplication. While the prosodic structure of the other patterns is also found in nonreduplicated forms (see (2)), the prosodic structure of $C_1C_2VC_3C_{2C}VC_{3C}$ is limited to reduplicated stems (with the exception of blends; e.g., *cfargol*, a hybrid of *cfardéa* 'frog' and *xargol* 'grasshopper'). It has been suggested that since an initial complex onset is found in suffixed forms (e.g., *gamal – gmalim* 'camel(s)'), this pattern involves suffixation. Further support can be drawn from Graf's (2002) observation that the types of stems that undergo deletion under suffixation (CéCVC and CaCVC) are also the ones that can serve as bases for $C_1C_2VC_3C_{2C}VC_{3C}$ reduplication; that is, from *kélev* 'dog' we can get *klavlav* 'little dog' (cf. *klavim* 'dogs'), but from *kipod* 'hedgehog' we cannot get **kpadpad/kpodpod* (cf. *kipodim* 'hedgehogs').

I claim that these structural properties are not sufficient to grant $C_1C_2VC_3C_{2C}VC_{3C}$ a different status and thus a different formal account. First, CaCVC stems that do not undergo deletion under suffixation (e.g., *katav – katavim* 'reporter(s)') are still potential bases for $C_1C_2VC_3C_{2C}VC_{3C}$ reduplication (*ktavtav* 'reporter (derogatory)'). Second, Hebrew nouns do not have consonant-initial suffixes, and therefore there are no suffixed forms of the structure CCVC-CVC.[18] Third, complex onsets are found in stems without a suffix (e.g., *cfardéa* 'frog', *ʃvil* 'path'). Thus, the prosodic structure of $C_1C_2VC_3C_{2C}VC_{3C}$ shares properties with both bare stems and suffixed forms, and therefore it cannot constitute an argument for either.

There is, however, a positive structural reason to view a $C_1C_2VC_3C_{2C}VC_{3C}$ pattern as a stem, rather than a suffixed form. As noted in Bat-El 1989, the vocalic patterns found in $C_1C_2VC_3C_{2C}VC_{3C}$ reduplicated forms also appear in nonreduplicated ones (as is the case with the other reduplicated forms; see (2)); for example, {aa} in *vradrad* 'pinkish' and *sarbal* 'overall

[18] One exception is *gvartan* 'valiant', derived from *géver* 'man' plus *-an*. Here, instead of **gavran*, we find *gvartan*; compare *kéver* 'grave' – *kavran* 'undertaker'.

(garment)', {au} in *ʔasafsuf* 'crowd' (epenthetic vowel in the first syllable; see footnote 3) and *yalkut* 'bag', and {ao} in *cmarmóret* 'shiver' and *masóret* 'tradition'.

The absence of coherent semantic distinctions between the different patterns of reduplication and a prosodic trigger for all cases of reduplication (see section 1) supports two aspects of the approach advocated in this article: (a) all patterns of reduplication should be analyzed under a unified formal account (section 4), and (b) reduplication is lexically specified, as much as other word formation strategies (section 1).

## 6 Conclusions

Identical consonants in Hebrew may appear anywhere in the stem. However, I have argued that speakers distinguish between stems where the identical consonants are at the right periphery and stems where the identical consonants are in other positions. This distinction is accessible online, without reference to a lexical base. I have proposed constraints that allow speakers to make this distinction and to assign a reduplicated structure only to stems with identical consonants at the right edge. To parse a stem as reduplicated means to identify the rightmost member of a pair of identical consonants as a copy of the other. When parsing a reduplicated stem, speakers distinguish between (a) the consonants belonging to the domain of the base and (b) the residual consonants, which are the copy.

I have proposed that reduplicated stems are lexically associated with a constraint COPY, since reduplication is often not predicted on independent grounds (prosodic or semantic). This approach is in the spirit of work by Russell (1995, 1999), who views morphological processes as constraints (see also Yip 1998, Adam and Bat-El 2000, Hammond 2000, Bat-El 2002, 2003a). The advantage of viewing reduplication as triggered by a constraint, rather than by an affix in the input, is that a constraint is available for online parsing without reference to an input.

COPY may reside in one of three designated positions in the ranking of the prosodic constraints defining the licit prosodic structures. Its position, along with the number of different consonants, determines the pattern of reduplication. Note that COPY does not add a prosodic structure or a specific segmental unit; rather, it requires every segment to have two occurrences in the stem. The copied material is thus a residue left after identifying a base. It has a status similar to that of the consonantal root in nonreduplicated stems, which, as argued in Bat-El 2001, 2003a,b, is not a morphological unit but a residue left after stripping away the configuration composed of the vocalic pattern and the prosodic structure. That is, words may consist of units directly referred to by the grammar plus residual segmental material.

This account differs from that of Gafos (1998), who proposes that the vocalic pattern is associated with RED, and thus every stem is potentially reduplicated; the fact that reduplication is not surface true in all stems is due to prosodic constraints. To account for the lexical distinction between the patterns $C_1VC_2VC_{2C}$ and $C_1VC_2C_{1C}VC_{2C}$, Gafos proposes that the latter is lexically associated with the template $[\sigma_{\mu\mu}\sigma]$.[19] There are several problems with this account. First, there

[19] The pattern $C_1VC_2C_3VC_{3C}$, which contrary to Gafos's prediction does exist, will also be lexically specified for $[\sigma_{\mu\mu}\sigma]$.

is no independent motivation for a moraic structure in Hebrew, since the language does not exhibit a weight distinction; there are no phonemic long vowels, nor is stress attracted by CVC syllables (Bolozky 1978a). McCarthy (1984) has shown that Hebrew templates must be based on syllables, without further specification, as the difference between verbs like *gidel* 'to raise', *tilfen* 'to phone', *tilgref* 'to telegraph', and *ʃnirkel* 'to use a snorkel' is not due to the internal structure of the syllables in the template. Moreover, the template $[\sigma_{\mu\mu}\sigma]$ cannot distinguish between the patterns $C_1C_2VC_3C_{2C}VC_{3C}$ and $C_1VC_2C_3VC_{3C}$, since there is no moraic distinction between simple and complex onsets. Second, Gafos's analysis wrongly predicts that the expansion of monosyllabic bases will always take place via reduplication, since RED is associated with the vocalic pattern. As shown in (5c), monosyllabic bases can be expanded by means other than reduplication, and thus a further lexical specification will be required to distinguish between forms like *ʃem* 'name' – *ʃiyem* 'to name' and *ken* 'nest' – *kinen* 'to nest'. Also, the absence of reduplication in pairs like *ʃir* 'song' – *ʃar* 'to sing' needs to be lexically specified.

Here, I argued that all reduplicated forms are lexically associated with COPY. There is nothing idiosyncratic in the patterns requiring, as Gafos proposes, the lexically specified template $[\sigma_{\mu\mu}\sigma]$ (i.e., $C_1VC_2C_3VC_{3C}$ and $C_1VC_2C_{1C}VC_{2C}$); therefore, a formal distinction does not reflect their status in the grammar, which is the same as that of the other patterns. The lexical specification suggested here is on a par with the specification required for other morphological processes (see end of section 1), as the presence of several word formation strategies for a particular function renders the selection of a particular strategy unpredictable in many cases.

Nonreduplicated stems with identical consonants are not banned, but they are certainly disfavored, as suggested by experimental studies (and their low type-frequency in the Hebrew lexicon). SCORRI, which bans such forms, is ranked low enough to allow them to be parsable and derivable; they are parsable because SCORRI is ranked below SCORRP, and they are derivable because SCORRI (which does the job of the OCP) is ranked below MAX. These rankings account for the fact that not only do forms with identical consonants in positions other than the right periphery of the stem exist in Hebrew, they are also not rejected by the grammar. An analysis that does not predict the existence of such forms does not reflect the native speaker's grammar.

## References

Adam, Galit. 2002. From variable to optimal grammar: Evidence from language acquisition and language change. Doctoral dissertation, Tel Aviv University.

Adam, Galit, and Outi Bat-El. 2000. Morphological knowledge without morphological structure: The emergence of inflectional suffixes in the acquisition of Hebrew verbs. Paper presented at the workshop ''Phonology and Prosodic Morphology of Semitic Languages,'' Tel Aviv. http://www.tau.ac.il/~obatel/.

Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language & Linguistic Theory* 20:1–42.

Bat-El, Outi. 1989. Phonology and word structure in Modern Hebrew. Doctoral dissertation, UCLA, Los Angeles, Calif.

Bat-El, Outi. 1994a. The optimal acronym word in Hebrew. In *Proceedings of the 1994 Annual Conference*

*of the Canadian Linguistic Association*, ed. by Päivi Koskinen, 23–37. Toronto Working Papers in Linguistics. Toronto: University of Toronto, Department of Linguistics, Linguistics Graduate Course Union.

Bat-El, Outi. 1994b. Stem modification and cluster transfer in Modern Hebrew. *Natural Language & Linguistic Theory* 12:571–596.

Bat-El, Outi. 1995. Resolving prosodic mismatch in Modern Hebrew verb formation. In *Leiden in last: HIL phonology papers I*, ed. by Harry van der Hulst and Jeroen van de Weijer, 25–40. The Hague: Holland Academic Graphics.

Bat-El, Outi. 1997. On the visibility of word internal morphological features. *Linguistics* 35:289–316.

Bat-El, Outi. 2001. In search of the roots of the C-root: The essence of Semitic morphology. Paper presented at the workshop ''Roots and Template Morphology,'' University of Southern California, Los Angeles. http://www.tau.ac.il/~obatel/.

Bat-El, Outi. 2002. True truncation in Hebrew imperatives. *Language* 78:1–33.

Bat-El, Outi. 2003a. The fate of the consonantal root and the binyan in Optimality Theory. *Recherches Linguistiques de Vincennes* 32:31–60.

Bat-El, Outi. 2003b. Semitic verb structure within a universal perspective. In *Language processing and acquisition in languages of Semitic, root-based, morphology*, ed. by Joseph Shimron, 29–59. Amsterdam: John Benjamins.

Berent, Iris, Daniel Everett, and Joseph Shimron. 2001. Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology* 42:1–60.

Berent, Iris, and Joseph Shimron. 1997. The representation of Hebrew words: Evidence from the Obligatory Contour Principle. *Cognition* 64:39–72.

Berent, Iris, and Joseph Shimron. 2003a. Co-occurrence restrictions on identical consonants in the Hebrew lexicon: Are they due to similarity? *Journal of Linguistics* 39:31–55.

Berent, Iris, and Joseph Shimron. 2003b. What is a root? In *Language processing and acquisition in languages of Semitic, root-based, morphology*, ed. by Joseph Shimron, 201–222. Amsterdam: John Benjamins.

Berent, Iris, Joseph Shimron, and Vered Vaknin. 2001. Phonological constraints on reading: Evidence from the Obligatory Contour Principle. *Journal of Memory and Language* 44:644–665.

Berhane, Girmay. 1992. Issues in the phonology and morphology of Tigrinya. Doctoral dissertation, Université du Québec à Montréal.

Bolozky, Shmuel. 1978a. Some aspects of Modern Hebrew phonology. A chapter in *Modern Hebrew structure* by Ruth A. Berman, 11–67. Tel Aviv: University Publishing Projects.

Bolozky, Shmuel. 1978b. Word formation strategies in the Hebrew verb system: Denominative verbs. *Afroasiatic Linguistics* 5:111–136.

Bolozky, Shmuel. 1999. *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: Brill.

Booij, Geert E. 1996. Inherent versus contextual inflection and the split morphology hypothesis. In *Yearbook of morphology 1995*, ed. by Geert E. Booij and Jaap van Marle, 1–16. Dordrecht: Kluwer.

Everett, Daniel, and Iris Berent. 1997. The comparative optimality of Hebrew roots: An experimental approach to violable identity constraints. Rutgers Optimality Archive ROA 235–1297. http://roa.rutgers.edu.

Frisch, Stefan, Janet Pierrehumbert, and Michael Broe. 2004. Similarity avoidance and the OCP. *Natural Language & Linguistic Theory* 22:179–228.

Gafos, Diamandis. 1998. Eliminating long-distance consonantal spreading. *Natural Language & Linguistic Theory* 16:223–278.

Graf, Dafna. 2002. A study on nominal reduplication in Modern Hebrew. In *Proceedings of the Israel Association for Theoretical Linguistics 18*, ed. by Yehuda N. Falk. Available at atar.mscc.huji.ac.il/~english/IATL/18/TOC.html.

Greenberg, Joseph. 1950. The patterning of root morphemes in Semitic. *Word* 6:162–181.
Hammond, Michael. 2000. There is no lexicon! In *Coyote papers: Working papers in linguistics 10*, ed. by Sean Hendricks, Sachiko Ohno, Amy Fountain, Debora Cole, and Mizuki Miyashita, 55–77. Tucson: University of Arizona Linguistics Circle.
Inkelas, Sharon. 1994. The consequence of optimization for underspecification. In *NELS 25*, ed. by Jill Beckman, 187–302. Amherst: University of Massachusetts, GLSA.
Itô, Junko, and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In *Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 181–209. University of Massachusetts Occasional Papers in Linguistics 18. Amherst: University of Massachusetts, GLSA.
Itô, Junko, Armin Mester, and Jaye Padgett. 1995. Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26:571–614.
Leslau, Wolf. 1967. *Amharic textbook*. Wiesbaden: Harrassowitz.
Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13:435–482.
McCarthy, John. 1979. Formal problems in Semitic phonology and morphology. Doctoral dissertation, MIT, Cambridge, Mass.
McCarthy, John. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12:373–418.
McCarthy, John. 1984. Prosodic organization in morphology. In *Language sound structure*, ed. by Mark Aronoff and Richard Oehrle, 299–317. Cambridge, Mass.: MIT Press.
McCarthy, John. 1993. A case of surface constraint violation. *Canadian Journal of Linguistics* 38:169–195.
McCarthy, John. 2004. Taking a free ride in morphophonemic learning. Rutgers Optimality Archive ROA 683-0904. http://roa.rutgers.edu.
McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. In *Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 249–384. University of Massachusetts Occasional Papers in Linguistics 18. Amherst: University of Massachusetts, GLSA.
McCarthy, John, and Alan Prince. 1999. Faithfulness and identity in Prosodic Morphology. In *The prosody-morphology interface*, ed. by René Kager, Harry van der Hulst, and Wim Zonneveld, 218–309. Cambridge: Cambridge University Press.
Pierrehumbert, Janet. 1993. Dissimilarity in the Arabic verbal roots. In *NELS 23*, ed. by Amy J. Schafer, 367–381. Amherst: University of Massachusetts, GLSA.
Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. Report RuCCS-TR-2, Center for Cognitive Science, Rutgers University, New Brunswick, N.J. Rutgers Optimality Archive ROA 537. http://roa.rutgers.edu.
Rose, Sharon. 2000. Rethinking geminates, long-distance geminates, and the OCP. *Linguistic Inquiry* 31: 85–122.
Rose, Sharon, and Rachel Walker. 2001. A typology of consonant agreement as correspondence. Rutgers Optimality Archive ROA 458-0801. http://roa.rutgers.edu.
Russell, Kevin. 1995. Morphemes and candidates in Optimality Theory. Rutgers Optimality Archive ROA 44-0195. http://roa.rutgers.edu.
Russell, Kevin. 1999. MOT: Sketch of an OT approach to morphology. Rutgers Optimality Archive ROA 352-1099. http://roa.rutgers.edu.
Schwarzwald, Ora. 1974. shorashim, bsisim u-mivne ha-morfemot (Roots, bases, and the structure of morphemes). *Leshonenu* 38:131–136.
Ussishkin, Adam. 1999. The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence. *Phonology* 16:401–442.
Ussishkin, Adam. 2000. The emergence of fixed prosody. Doctoral dissertation, University of California at Santa Cruz.
Yip, Moira. 1996. Feet, tonal reduction and speech rate at the word and phrase level in Chinese. Rutgers Optimality Archive ROA 159-0896. http://roa.rutgers.edu.

Yip, Moira. 1998. Identity avoidance in phonology and morphology. In *Morphology and its relation to phonology and syntax*, ed. by Steven G. Lapointe, Diane Brentari, and Patrick M. Farrell, 216–246. Stanford, Calif.: CSLI Publications.

Zadok, Gila. 2002. Abbreviations: A unified analysis of acronym words, clippings, clipped compounds, and hypocoristics. Master's thesis, Tel Aviv University.

*Department of Linguistics*
*Tel Aviv University*
*Tel Aviv 69978*
*Israel*

*obatel@post.tau.ac.il*