THE HEBREW LANGUAGE IN ISRAEL

2. PARSING FORMS WITH IDENTICAL CONSONANTS: HEBREW REDUPLICATION

OUTI BAT-EL
Tel-Aviv University

0. INTRODUCTION

Au: Check Sec. no.

Native speakers' knowledge of the morphology of their language is manifested by their ability to identify structural relations between words and to coin new words on the basis of existing ones. In addition, speakers are able to assign a structure, i.e. to parse a word, without reference to a related word. This capacity is manifested by parsing "orphan" complex words (i.e. complex words that do not have related non-complex counterparts), as well as nonsense words provided in experiments. In this paper, I draw attention to the latter capacity, i.e. the morphological parsing of words without reference to a related word.

The morphological construction under consideration here is reduplication, identified by one or two pairs of identical consonants on the right periphery of the stem (*xagag* 'to celebrate' *ʃixrer* 'to release', *kilkel* 'to spoil', *klavlav* 'little dog'). Identical consonants can appear also on the right periphery of the stem (cf. *mimen* 'to finance' vs. *minen* 'to aportion'), but only those with identical consonants on the right periphery are considered reduplicated by native speakers. To account for this right-edge preference, I propose surface ranked constraints that allow speakers to parse all and only such forms as reduplicated.

1. HEBREW REDUPLICATION

In the description of reduplication below I show that the structure of reduplicated forms is limited by the same prosodic and segmental restrictions imposed on

non-reduplicated forms (§1.1). I also claim that limitation of reduplication to the right periphery of the stem is contingent upon the morphology of affixation as well as general constraints on processing (§1.2).

1.1. The form and meaning of reduplication

Hebrew words may consist of a bare stem (*gadal* 'to grow', *gódel* 'size'), or a stem plus affixes (*h-igdil* 'to enlarge', *m-igdal* 'tower', *migdal-on* 'little tower'). Structural relations among words are expressed by apophony (i.e. alternation in vocalic pattern, also called ablaut), affixation, and alternation in prosodic structure, where the latter is often accompanied by affixation.

Reduplicated forms, which can be nouns, adjectives, or verbs, are in most cases structurally identical to non-reduplicated ones, in terms of vocalic pattern, prosodic structure, and affixation (cf. *cided* 'to side with' vs. *gidel* 'to raise', *xalil* 'flute' vs. *sakin* 'knife', *m-ixlal-a* 'college' vs. *m-iʃtar-a* 'police'). Reduplicated stems, like most non-reduplicated ones, are disyllabic, where a complex syllable margin (consisting of two consonants) is allowed only in word initial position (in nouns). Within these prosodic limitations on the number and structure of syllables, reduplicated forms can have one or two pairs of identical consonants: one pair in the schematic pattern $C_hV(C_i)C_jVC_j$ (*cided* 'to side with', *xamcic* 'sour grass') and two pairs in the schematic pattern $(C_h)C_iVC_jC_iVC_j$ (*gilgel* 'to roll', *vradrad* 'pinkish'). Three pairs of identical consonants cannot fit into a disyllabic stem due to the impermissible medial complex onset (*$C_hC_iVC_j.C_hC_jVC_j$). Due to space restrictions, I limit the discussion here to forms with one pair of identical consonants (see Bat-El 2002a for an extended discussion).

Studies of Hebrew reduplication, in particular in denominative verbs (Bat-El 1994, Gafos 1989, Ussishkin 2000) suggest that reduplication is triggered by the requirement to form a disyllabic stem (*cad* 'side' - *cided* 'to side with'). However, there are also reduplicated forms whose non-reduplicated counterpart is already disyllabic (*dover* 'spokesman' - *divrer* 'to speak as a spokesman', */iʃer* 'to confirm' - */iʃrer* 'to reconfirm'). Thus, we cannot claim that reduplication is always triggered by prosodic restrictions.

Semantically, many reduplicated forms are associated with a diminutive property for nouns (*dagig* 'little fish') and, as argued in Ussishkin (2000), durative/repetitive for verbs (*kidrer* 'to dribble'). However, we cannot claim that reduplication is triggered by these semantic properties, since there are quite a few reduplicated forms that do not carry these semantic properties (*gal* 'wave' - *galil* 'cylinder', *ʃakal* 'to weight' - *ʃiklel* 'to consider relative weight'). We may attribute the semantic properties found in quite a few reduplicated forms to what Bolozky (1978) calls "sporadic analogies". Bolozky gives examples like *himxiz* 'to make (a novel) into a play' (derived from *maxaze* 'play'), which adopts the structure of the semantically related verbs *hisrit* 'to make/shoot/project a movie' (from *séret* 'movie') and *hilxin* 'to compose a tune' (from *láxan* 'tune'). However, whatever the semantic property shared by these verbs is, it cannot be attributed to the structure, which hosts verbs of a garden variety of semantic

P1: MRM/SJS P2: MRM/SJS QC: MRM/SJS T1: MRM
KI076-02 May 25, 2004 2:13 Char Count= 0

properties. The same is true for the general *-on* suffix, which appears in many different nouns (*ʃabat-on* 'sabbatical' */avir-on* 'airplane', *halix-on* 'treadmill'), but also in most nouns related to newspaper (*/it-on* 'newspaper', *mekom-on* 'local newspaper', *ʃvu/-on* 'weekly newspaper', *yarx-on* 'monthly newspaper'). Here again, one cannot claim that the suffix *-on* assigns the meaning 'disposable reading material', as it appears in plenty of forms that do not carry this meaning.

I thus claim that reduplication is a strategy of stem formation not necessarily associated with a specific meaning or triggered by prosodic constraints. Its general purpose is to form a different but related word, as is the case with other cases of templatic word formation in Hebrew (*ʃémen* 'oil' - *ʃamen* 'fat' - *ʃuman* 'grease' - *ʃamén-et* 'cream').

1.2. The right edge preference

Reduplication in Hebrew is restricted to the right periphery of the stem; that is, while *simem* 'to drug' is reduplicated, *mimen* 'to finance' is not. I suggest that this right edge preference is in part due to affixation, in particular to the distinction between template-internal prefixes and template-external suffixes. Most Hebrew prefixes are internal to the disyllabic template of the stem (*h-igdil* 'to enlarge', *m-igdal* 'tower', *y-igdal* 'he will grow'), while suffixes are usually external (*higdíl-u* 'they enlarged', *migdal-im* 'towers', *migdal-on* 'little tower').[1]

When a vowel initial suffix is added to a CVCVC verb stem, a non-high vowel in the stem final syllable is deleted (*gadal-a* ---> *gadla* 'she grew', *gidel-u* ---> *gidlu* 'they raised'). In nouns and adjectives, only *e* is deleted in this position (*makel-ot* ---> *maklot* 'sticks', *xiver-im* ---> *xivrim* 'pale pl.'). In reduplicated forms with such prosodic structure, the vowel in the stem final syllable is replaced by *e* to rescue a violation of the Obligatory Contour Principle (OCP), which does not allow adjacent identical segments (*garar-a* ---> *garer-a* (**garr-a*) 'she dragged'). Crucially, the *e* in *garer-a* does not demolish the template of the stem *garar*, as it occupies a position of a stem vowel.

This is, however, not the case when the two identical consonants are on the left periphery. If we take a verb with two identical consonants on the left periphery, like *mimen*, and try to fit it into binyan hiCCiC (as in *h-igdil*), we would get **h-immin*. To rescue the OCP violation encountered by the two adjacent *m*'s we have to insert an *e*. The epenthetic *e* would demolish the template (**h-imemin*), as it requires a new position not presented in the template of this binyan. The same goes for the formation of a taCCiC (as in *t-almid*) type noun from *mimen* (**t-ammin*, **t-amemin*), or any form with an internal prefix.[2]

$C_iVC_iVC_j$ stems, i.e. stems with two identical consonants on the left periphery, are thus derivationally limited, and therefore, we do not expect a language with

[1] Some affixes can be both internal and external. For example, the participle *m-*, which is internal in *m-agdil* 'he is enlarging' and external in *me-gadel* 'he is raising', and the agent suffix *-an*, which is internal in *rakda-n* 'dancer' and external in *taklit-an* 'DJ'. See Bolozky and Schwarzwald (1992) for this distinction with respect to the suffix *-ut*.

[2] I know of only one case where the template is demolished, *xacocra* 'trumpet'-*xicecer* 'to play the trumpet' (cf. *tilfen*). However, *xicecer* has an alternative, more common form, *xicrec*.

P1: MRM/SJS P2: MRM/SJS QC: MRM/SJS T1: MRM
KI076-02 May 25, 2004 2:13 Char Count= 0

template-internal prefixes to have reduplication on the left periphery.[3] The priority of the prosodic restrictions does not allow the learner to generate left-edge reduplication.

This type of paradigmatic limitation is not restricted to reduplication. For example, the binyan pattern CaCaC, as in *gadal* 'to grow', has more verbs in the Hebrew lexicon than any other binyan (Berman 1989), but the binyan pattern CiCeC, as in *gidel* 'to raise', is more productive, i.e. it is often selected for denominative verbs (the most common way for creating new verbs in Hebrew). The reason is that many nouns that serve as the base for denominative verbs have four or more consonants, and the future form of CaCaC verbs cannot accept more than three consonants, due to its prosodic alternation in the inflectional paradigm and the prohibition against medial complex onsets.[4] Thus, the hypothetical verb **maspar* (cf. *gadal*) would not have a future form, since **y-im.spar* (cf. *y-igdal*) has a medial complex onset, and **y-imsepar* has an epenthetic vowel which obscures the template. In CiCeC verbs, there is no prosodic alternation in the inflectional paradigm (*gidel* - *ye-gadel* 'to raise Past - Future') and therefore, this pattern can host any type of denominative verbs (as long as the complex onsets and codas do not violate the Sonority Sequencing Generalization or the OCP).

Beyond the language internal explanation, there is also a psycholinguistic reasoning for right edge reduplication. Models of word recognition, such as in Marslen-Wilson (1987) suggest that the lexicon is activated by the first (i.e. the leftmost) one or two segments in the input word (i.e. the word the speaker encounters) and proceeds from there until only one candidate consistent with the input remains. Thus, as suggested in Bat-El (to appear), recognition of right-edge reduplicated form $\{C_iVC_jV\}C_j$, where the base (enclosed in $\{\}$) is on the left edge, would be faster than recognition of left-edge reduplicated form $C_i\{VC_jVC_j\}$, where the base is further to the right.

This view is supported by the findings reported in Berman (1990), that children often add a consonant to a derived word in order to arrive at an optimal prosodic structure. The site of the new consonant, whether a copy of an existing consonant or an independent coronal consonant, is usually at the right edge of the word (*kise* 'chair' ---> *me-kase**s**, me-kase**t**, me-kase**n***).

Steriade (1993) and Beckman (1997) argue that some positions in the word, including word initial position (and stressed syllables), are psycholinguistically more salient, and can thus be targeted by the grammar. Such positions, as claimed in Nelson (1998), are the target of affixal reduplication, and therefore, there is preference for left-edge affixal reduplication (or reduplication of stressed syllable), as in Ponapean *duu-duupek* 'starved Durative'. However, Hebrew reduplication is not affixal but rather prosodic, as the copied material resides within the prosodic template of the stem, which, as claimed in §1.1, is the same prosodic template found in non-reduplicated forms. Nelson (2002) argues, on the basis of Yoruba emphatic ideophones (*rogodo-do* 'of being very round and small'), that unlike affixal reduplication, prosodic reduplication prefers the right-edge.

[3] Indeed, the nominal $C_iVC_iVC_j$ forms are without a prefix (e.g. *nanas* 'dwarf'), and the verbal $C_iVC_iVC_j$ forms are in a binyan whose prefix is external (*mimen*, *me-mamen*, *ye-mamen* 'to finance Past, Participle, Future').
[4] A medial complex onset is permissible only when it appears throughout the inflectional paradigm (*til.gref—ye-.tal.gref* 'to telegraph Past—Future').

P1: MRM/SJS P2: MRM/SJS QC: MRM/SJS T1: MRM
KI076-02 May 25, 2004 2:13 Char Count= 0

This proposal is consistent with the psycholinguistic explanation given above for the prosodic right-edge reduplication in Hebrew.

2. THE *minen - mimen* PROBLEM

One of the central issues in the study of Hebrew reduplication (and Semitic in general) has been concerned with the **sasam - samam* legacy, according to which identical consonants appear only on the right periphery of the stem, i.e. there are no forms like *sasam* (see Greenberg's 1950 co-occurrence restrictions). McCarthy (1979, 1981) attributes this fact to the OCP, which prohibits adjacent identical root consonants, and thus excludes both {ssm} or {smm} as possible roots; the only possible root is {sm}.

McCarthy assumes a non-linear representation, where distinct segmental morphemes (the consonantal root and the vocalic pattern, as well as affixes) appear on different tiers and associated with a prosodic template. The root consonants are associated with the prosodic template in one-to-one left-to-right fashion. Since there are three consonantal positions in a CVCVC template, but only two consonants in the root {sm}, one of the consonants spreads to the third position. Since association begins at the left edge, it is always the rightmost consonant that occupies two positions (the same goes for a CVCCVC template and a triconsonantal root). Under this approach, the language can generate $C_hV(C_i)C_jVC_j$ forms, but not $C_iVC_iVC_j$ and $C_hVC_hC_iVC_j$ forms.

However, Hebrew does have $C_iVC_iVC_j$ and $C_hVC_hC_iVC_j$ forms, such as *sisgen* 'to variegate', as well as *mimen* 'to finance', which minimally contrasts with *minen* 'to apportion', and *mimeʃ* 'to materialize' which contrasts with *miʃeʃ* 'to feel with the hands'. There are not many forms of this type, but we cannot treat them as sporadic exceptions because they can freely enter the language via borrowing. The most recent example, to the best of my knowledge, is the verb *titel* 'to diaper', which has been formed from a brand name for diapers *titulim*, exactly like *mimen* 'to finance' from the Greek loan word *mamon* 'finance' (see Bolozky 1978 and Bat-El 1994 for the formation of verbs from nouns and adjectives).

There is, however, a major difference between $C_iVC_jVC_j$ (*minen*) and $C_iVC_iVC_j$ (*mimen*) forms; $C_iVC_jVC_j$ forms are reduplicated while $C_iVC_iVC_j$ are not. This is supported by fact that there is no form with identical consonants on the left periphery (i.e. $C_iVC_iVC_j$ or $C_hVC_hC_iVC_j$) that has a related form with one occurrence of the consonants (i.e. C_iVC_j or $C_hVC_iVC_j$ respectively). That is, while identical consonants can appear at either edge of the stem, only forms with identical consonants at the right edge are reduplicated. Within McCarthy's (1981) approach, this means that there are roots like {ssm} but not like {smm}. That is, the consonantal root can violate the OCP, but only at the left edge.[5]

[5] This problem can be eliminated if we assume that the OCP holds for primary roots only, and that some nouns cannot be decomposed to a root-plus-binyan. Under this assumption, *mamon* does not have a root, and the root of *mimen*, which is derived from *mamon*, is secondary. This solution is, however, local. As lexical items can vanish through history (as happened to some of the bases of some orphan reduplicated forms), the possible disappearance of *mamon* from the vocabulary of (a future stage of) Hebrew would leave the verb *mimen* with a primary root {mmn}.

P1: MRM/SJS P2: MRM/SJS QC: MRM/SJS T1: MRM
KI076-02 May 25, 2004 2:13 Char Count= 0

Notice also that there are forms with identical consonants at the right edge that are derived from bases with identical consonants. For example, *bises* 'to establish', like *mimen*, is derived from the Greek loan word *basis* 'base'. This may suggest that *basis*, and thus *bises*, are not reduplicated. That is, that some $C_iVC_jVC_j$ forms are reduplicated (*cided*, from *cad*) and others are not (*bises* from *basis*). Also, orphan forms with identical consonants (*garar* 'to tow') could then be considered as non-reduplicated since they do not have a related non-reduplicated counterpart.

I claim that a grammar that allows such structural ambiguity, i.e. that some $C_iVC_jVC_j$ forms are reduplicated and others are not, is unlearnable and unprocessable. Speakers overgeneralize the structure obtained from surface morphological relations to all forms with the relevant identical structure, in this case identical consonants on the right periphery of the stem. Given the vast amount of non-reduplicated—reduplicated pairs in the language, every $C_hV(C_i)C_jVC_j$ is analyzed by speakers as reduplicated. That is, a systematic structural ambiguity (i.e. not a sporadic homophony), without a semantic or syntactic correlation, is unlearnable. Moreover, speakers would not be able to determine whether a form with two identical consonants on the right periphery is reduplicated or non-reduplicated in a morphological parsing task, i.e. without reference to a base. We could posit an abstract underlying representation for the orphan reduplicated forms, but we still have to account for speakers' ability to identify the morphological structure of nonsense words, for which abstract underlying representation is not available.

The claim that all $C_iVC_jVC_j$ forms are parsed as reduplicated gains support from experimental studies. Berent and Shimron (1997) conducted two experiments, which examined the acceptability of nonsense words with two identical consonants on the left periphery (*siseg*), on the right periphery (*sigeg*), and without identical consonants (*riseg*). One experiment involved relative rating (of triplets), and the other absolute rating (of a random list). In the relative rating, there is a significant preference *riseg* > *sigeg* > *siseg*, which suggests that when given the option, non-reduplicated forms are preferred. In the absolute rating, there is no significant difference of acceptability between *sigeg* and *riseg*, but *siseg* is still rated low; *riseg, sigeg* > *siseg*. The low acceptability of *siseg* reflects the effect of the OCP. The similar rating of *riseg* and *sigeg* (in the absolute rating), and the preference of *sigeg* over *siseg*, suggest that *sigeg* is parsed as a reduplicated form.

Proponents of the similarity-based model of morphology, most popular among psycholinguists, may argue that nonsense and orphan $C_iVC_jVC_j$ forms are parsed as reduplicated in analogous to existing $C_iVC_jVC_j$ forms that have a base. As Hahn and Chater (1998) emphasize, the distinction between similarity-based and rule-based models has been rather intuitive and general. One of the intuitive criteria is that similarity is some function of common properties. Berent et al. (1999), in their study of nominal Hebrew plurals, take distinctive features as the common property for similarity. For example, *vinon* is similar to *vilon* 'curtain', as the two differ in the value of one feature in one segment (nasality in *n* vs. *l*); the similarity of *vikon* and *vilon* is less than that of *vilon* and *vinon*, since *vikon* and *vilon* differ in the value of two features in one segment (sonority and place of articulation in *k* vs. *l*); *kixon* is, however, dissimilar from

vilon as the two differ in at least two features in two segments (continuancy and place of articulation in *k* vs. *v* and sonority and place of articulation in *x* vs. *l*). Unfortunately, such a featural basis of similarity cannot tell us whether the nonsense word *sigeg* is more similar to *zikek* 'to purify' or *sigel* 'to adjust', as in the former, there is voicing distinction in three segments (*s* vs. *z* and the two *g*'s vs. the two *k*'s), and in the latter, two features in one segment (place of articulation and sonority in *g* vs. *l*). Similarly, it cannot tell us whether the orphan reduplicated form *salal* 'to pave' is more similar to the unreduplicated form *salad* 'to spring back' or to the reduplicated form *sarar* 'to be stubborn'. Moreover, as shown in Berent et al. (2002), speakers interpret nonsense $C_iVC_jVC_j$ forms with non-native segments as reduplicated, without a basis for featural similarity. Berent et al. thus argue that "a complete account of linguistic processing must incorporate mechanisms for generalization outside the representational space of trained items" (p. 113).

In what follows, I propose a mechanism based on ranked constraints, that allows speakers to parse all and only forms with identical consonants on the right periphery as reduplicated; forms with identical consonants in other sites are parsed as non-reduplicated.

4. SURFACE CONSTRAINTS ON HEBREW REDUPLICATION

The analysis below adopts the spirit (but refrains from the formal aspects) of the framework of Optimality Theory (Prince and Smolensky 1993).[6] Since Optimality Theory is an output-oriented framework, it is also suitable to account for parsing, where a structure has to be assigned to a surface form. One of the theory's components is a set of violable constraints, which are ranked on language specific grounds. The constraints often compete with each other in a given form, but as they are violable, the lower-ranked one has to be violated in order for the higher-ranked one to be satisfied. This framework thus allows maintaining a constraint like the OCP in the grammar, despite the presence of forms that violate it. However, in order for the OCP to be violated, it must be ranked below a competing constraint, which forces its violation.

A surface form has two morphological domains, a base and a stem, where the base is nested within the stem, $[\ \ldots\ {}_{B}\square\square\square\ldots]_{Stem}$ (in non-reduplicated stems the base and the stem overlap; $[\ \ldots\ {}_{B}\square\square s]_{Stem}$). Crucially, the copied material must reside outside the base, and the base must consist of all and only base segments (epenthesis ignored).[7] Thus, $C_1VC_{2C}V\ C_2]$ (where the subscript C indicates a copied segment) is not a licit structure since a base segment (C_2) is outside the base. Similarly, $C_1VC_2VC_{2C}\ ]$ is not a licit structure since a non-base segment (C_{2C}) is within the base. That is, the domain restriction provides the distinction between the base segments and the copied ones, and determines which of the identical consonants is the copy. I view the domain

[6] For a more detailed presentation of Optimality Theory see Archangeli and Langendoen (1997), Kager (1999), McCarthy (2002), and articles in Rutgers Optimality Archive (ROA) <http://ruccs.rutgers.edu/roa.html>.

[7] As argued in §1.1, there are cases where reduplication cannot be independently motivated by the prosodic or semantic properties. I thus assume that reduplicated forms are lexically associated with the morphological constraint Copy (see Yip 1995, Russel 1995, 1999, Adam and Bat-El 2000, Adam 2002, and Bat-El 2002b for a constraint-based approach to morphology). I ignore Copy in the discussion here as the issue here is parsing rather than derivation.

P1: MRM/SJS P2: MRM/SJS QC: MRM/SJS T1: MRM
KI076-02 May 25, 2004 2:13 Char Count= 0

restriction as an undominated constraint in Hebrew, and thus never violated (as there is no higher-ranked constraint that can force its violation).

I assume a constraint called SURFACE CORRESPONDENCE BY IDENTITY, which assigns a correspondence relation to any two identical consonants in a stem, i.e. regardless of their position. This constraint is respected when any two identical consonants in a surface form are parsed as a reduplicated pair, such that one of them is a base segment and the other is a copied segment. This constraint thus prohibits identical consonants in a base. The prohibition against identical (not necessarily adjacent) consonants in a base reduces, in a systematic way, the number of possible bases; a language does not need as many contrasting bases as provided by all the possible permutations and repetitions of the consonants. Indeed, the co-occurrence restrictions on the base consonants found in Semitic languages (Greenberg 1950) show that identical consonants across another consonant ($C_iVC_jVC_i$) are also disfavored, though not as much as identical adjacent consonants ($C_iVC_iVC_j$).[8] Thus, SURFACE CORRESPONDENCE BY IDENTITY does the job of the OCP in earlier analyses (see §2), but it makes a stronger claim since it does not require adjacency. It is compatible with Yip's (1998) definition of the OCP, which states that "output must not contain two identical elements" (p. 221), as well as Everett and Berent's (1997) constraint *IDENTITY, which does not allow two identical consonants in a root.[9]

As claimed earlier, $C_iVC_iVC_j$ forms are not reduplicated, and thus the identical consonants are not in correspondence relation, in violation of SURFACE CORRESPONDENCE BY IDENTITY. Therefore, it is necessary to appeal to a constraint that forces the violation of SURFACE CORRESPONDENCE BY IDENTITY. I propose the constraint SURFACE CORRESPONDENCE BY POSITION, which requires corresponding consonants to appear at the right edge of a domain, i.e. one at the right edge of the stem and the other at the right edge of the base. This constraint, supported by paradigmatic and recognition considerations (§1.2), allows only identical consonants on the right periphery of the stem to be in correspondence relations, i.e. one is the copy of the other.

Reduplicated forms, $C_hV(C_i)C_jVC_j$, respect both constraints. They are parsed as $[\{C_hV(C_i)C_jV\}C_{jC}]$, where one of the corresponding identical segments is at the right edge of the base (marked with }) and the other one, the copy, at the right edge of the stem (marked with]). Vowels are ignored as their position is subject to independent prosodic constraints.

The competition between the two constraints is manifested in forms with identical consonants in other sites. The fact that $C_iVC_iVC_j$ forms are not reduplicated, i.e. that the two identical consonants are base consonants, suggests that the constraint SURFACE CORRESPONDENCE BY POSITION has priority over SURFACE CORRESPONDENCE BY IDENTITY. That is, since $C_iVC_iVC_j$ (*mimen*) cannot be parsed as $[C_{iC}\{VC_iVC_j\}]$ due to SURFACE CORRESPONDENCE BY POSITION (as the identical consonants are at the

[8] See Pierrhumbert (1993) and Frisch (1996) for a gradient account of consonant cooccurrence restrictions in Arabic, based on degree of similarity and distance.

[9] The OCP in its original version has to be maintained for local adjacency, i.e. two surface adjacent identical consonants, as Hebrew has forms that violate SURFACE CORRESPONDENCE BY IDENTITY (*mimen*), but not forms that violate the OCP (**himmin*).

left, rather than the right edges of the domains), it must be parsed as [{$C_iVC_iVC_j$}] in violation of SURFACE CORRESPONDENCE BY IDENTITY. The same is true for forms with identical consonants in other sites, for example $C_hVC_iC_hVC_j$ (*safsal* 'bench'). In the structure [C_h{$VC_iC_hVC_j$}], one of the identical consonants is at the right, rather than the left edge of the stem, and the other is not at either edge. Thus, SURFACE CORRESPONDENCE BY POSITION forces the non-reduplicated parsing [{$C_hVC_iC_hVC_j$}], which violates SURFACE CORRESPONDENCE BY IDENTITY.

5. CONCLUSION

Hebrew speakers can parse forms with identical consonants on the right periphery as reduplicated, as well as forms with identical consonants in other positions as non-reduplicated, without reference to a base form. It was argued that this parsing is not based on similarity to reduplicated forms that have a base (i.e. analogy), but rather on surface constraints; one that requires correspondence between identical segments (SURFACE CORRESPONDENCE BY IDENTITY) and a higher-ranked one that restricts correspondence to the right periphery (SURFACE CORRESPONDENCE BY POSITION). Both constraints are also grounded outside the grammar, i.e. they have a functional motivation. SURFACE CORRESPONDENCE BY IDENTITY systematically reduces the number of bases in the lexicon, and SURFACE CORRESPONDENCE BY POSITION facilitates processing. Given these two ranked constraints, and the theoretical assumption that constraints are ranked and violable under competition, all forms with identical consonants on the right periphery are parsed as reduplicated regardless of their source. One vivid example is the recently coined blend *ʃiltet* 'to zap', which was formed as a blend from *ʃalat* 'remote control' plus *ʃotet* 'to wander around'; speakers who are not aware of this source (and there are quite a few) parse *ʃiltet* as the reduplicated counterpart of *ʃalat*.

REFERENCES

Adam, Galit. 2002. *From Variable to Optimal Grammar: Evidence from Language Acquisition and Language Change*. Ph.D. 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. A paper presented in a workshop on Phonology and Prosodic Morphology of Semitic Languages. Tel-Aviv. <http://www.tau.ac.il/humanities/lingui/people/batel. html>.

Arachangeli, Diana and Terence Langendoen (eds) 1997. *Optimality Theory: An Overview*. Oxford: Blackwell.

Bat-El, Outi. 1994. Stem modification and cluster transfer in Modern Hebrew. *NLLT* 12:571–596.

Bat-El, Outi. 2002a. Hebrew reduplication: The Interpretation of forms with identical consonants. A paper presented at UC Santa Cruz colloquium. <http://www.tau.ac.il/humanities/ lingui/people/batel.html>.

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

Bat-El, Outi. to appear. Semitic verb structure within a universal perspective. In J. Shimron (ed.) *Language, Processing, and Language acquisition in a Root-based Morphology*. Amsterdam: John Benjamins. pp. 29–59.

Beckman, Jill. 1997. Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* 14:1–46.

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

Berent, Iris, Steven Pinker, and Joseph Shimron. 1999. Default nominal inflection in Hebrew: evidence for mental variables. *Cognition* 72:1–44.

Berent, Iris, Gary Marcus, Joseph Shimron and Diamandis Gafos. 2002. The scope of linguistic generalization: Evidence from Hebrew word formation. *Cognition* 83:113–139.
Berman, Ruth. 1989. Children's knowledge of verb structure: Data from Hebrew. A paper presented at the 14th Boston University Conference on Language Development.
Berman, Ruth. 1989. How a new root is formed: Denominative verbs in Modern Hebrew. A paper presented at the 4th annual Meeting of the Israeli Association for Theoretical Linguistics. Jerusalem.
Bolozky, Shmuel. 1978. Word formation strategies in the Hebrew verb system: Denominative verbs. *Afroasiatic Linguistics* 5:111–136.
Bolozky, Shmuel and Ora Schwarzwald. 1992. On the derivation of Hebrew forms with the *+ut* suffix. *Hebrew Studies* 33:51–69.
Hahn, Ulrike and Nick Chater. 1998. Similarity and Rules: distinct? exhaustive? empirically distinguishable? *Cognition* 65:197–230.
Everett, Daniel and Iris Berent. 1997. The comparative optimality of Hebrew roots: An experimental approach to violable identity constraints. ROA 235-1297. <http://ruccs. rutgers.edu/roa.html>.
Frisch, Stefan. 1997. *Similarity and Frequency in Phonology*. Ph.D. dissertation, Northwest University. ROA 198-0597 (Rutgers Optimality Archive) <http://ruccs. rutgers.edu/roa.html>.
Gafos, Diamandis. 1998. Eliminating long-distance consonantal spreading. *NLLT* 16:223–278.
Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
Marsen-Wilson, W. 1987. Functional parallelism in spoken word recognition. *Cognition* 25:71–102.
McCarthy, John. 1979. *Formal Problems in Semitic Phonology and Morphology*. Ph.D. dissertation, MIT. [1982: IULC]
McCarthy, John. 1981. A prosodic theory of nonconcatenative morphology. *LI* 12:373–418.
McCarthy, John. 2002. *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.
Nelson, Nicole. 1998. Right Anchor, aweigh. ROA 284-0998. <http://ruccs.rutgers. edu/roa.html>.
Nelson, Nicole. 2002. Yoruba ideophones: Arguments for a "prosodic template" analysis. A paper presented in The Undergraduate Research Conference in UC Santa Cruz.
Pierrehumbert. J. 1993. Dissimilarity in the Arabic verbal root. *Proceedings of NELS 23*. Amherst: GLSA Publications.
Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. ms. Rutgers University and University of Colorado.
Russell, Kevin. 1995. Morphemes and candidates in Optimality Theory. ROA 44-0195. <http://ruccs.rutgers. edu/roa.html>.
Russell, Kevin. 1999. MOT: Sketch of an OT approach to morphology. ROA 352-1099. <http://ruccs.rutgers. edu/roa.html>.
Steriade, Donca. 1993. Positional neutralization. A paper presented at the 24th Meeting of the North East Linguistic Society. Amherst.
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*. Ph.D. dissertation, UCSC.
Yip, Moira. 1998. Identity avoidance in phonology and morphology. In S.G. Lapointe, B.K. Brentari, and P.M. Farrell (eds) *Morphology and Its Relation to Phonology and Syntax*. Stanford: CSLI Publications. pp. 216–246.