The representation of Hebrew words: Evidence from the obligatory contour principle

Iris Berent[a],*, Joseph Shimron[b]

[a]*Department of Psychology, Florida Atlantic University, 777 Glades Road, Boca Raton FL 33431-0991, USA*
[b]*School of Education, University of Haifa, Haifa 31905, Israel*

Abstract

The Hebrew root morpheme typically consists of three consonants. Hebrew allows a gemi-nation of a root consonant, but constrains its location [McCarthy, J. (1979). Formal problems in semitic phonology and morphology. Cambridge, MA; MIT Ph.D. dissertation. Distributed by Indiana University Linguistics Club. Garland Press, New York, 1985]. A gemination of a root-consonant is permitted at the end of the root (e.g., [mss]), but not at its beginning (e.g., [ssm]). Two experiments examined readers' sensitivity to the structure of the root morpheme by obtaining ratings for nonwords derived from nonroots. Root-initial gemination (e.g., [ssm]) was judged unacceptable compared to root-final gemination (e.g., [mss]) or no gemination controls (e.g., [psm]). The sensitivity to root structure emerged regardless of the position of the root in the word. These results have several implications. (1) Our findings demonstrate morphological decomposition. Hebrew speakers' ratings reflect a phonological constraint on the location of geminates. Being the domain of this constraint, the root morpheme must form a separate constituent in the representation of Hebrew words. (2) The rejection of root-initial gemination supports the psychological reality of the Obligatory Contour Principle, a pivotal constraint in autosegmental phonology. (3) A sensitivity to the location of geminates pre-supposes a distinction between the representation of geminate and nongeminate bigrams. Such a distinction, however, requires the implementation of a symbol. Our findings converge with numerous linguistic evidence in suggesting that the representation of constituency struc-ture is necessary to account for linguistic generalizations. © 1997 Elsevier Science B.V.

1. Introduction

There is a considerable body of evidence demonstrating the sensitivity of speakers of a variety of languages to the morphological structure of their native language

* Corresponding author. e-mail: iberent@fau.edu

(e.g., Caramazza et al., 1988; Emmorey, 1995; Feldman, 1994; Fowler et al., 1985; Grainger et al., 1991; Gordon, 1989; Kim et al., 1991; Marcus et al., 1995; Marslen-Wilson et al., 1994). However, there is an ongoing debate regarding the representational properties that must be postulated in order to account for this sensitivity. At the heart of the debate is the question of whether an account for speakers' morphological knowledge requires an explicit decomposition of words onto their formal constituents (e.g., Fodor and Pylyshyn, 1988; Pinker, 1991; Pinker and Prince, 1988), or, rather could this knowledge be explained without recourse to explicit formal constraints on word structure, in fact, without postulating explicit lexical entries at all (Rumelhart and McClelland, 1986; Seidenberg, 1987; Seidenberg and McClelland, 1989). On this latter view, morphological knowledge could be fully explained in terms of the co-occurrence of phonological, orthographic, and semantic properties of morphological units without assuming any formal morphological distinctions per sé. Indeed, most of the existing evidence for morphological compositionality comes from languages whose morphology is concatenative. In those languages, morphologically complex words may be formed by affixation without interrupting the integrity of the base morpheme. Consequently, morphemes tend to correspond to discrete orthographic and phonological units that are often associated with some well defined semantic features. An empirical dissociation of the contribution of formal units, morphemes, from that of their nonmorphological correlates is not easily achieved (Stemberger, 1995).

In the present research, we examine evidence for morphological compositionality in Hebrew, a language whose morphology is nonconcatenative. In this language, distinct morphemes are often interwoven, temporally co-occurring units. The root morpheme is not a linearly discrete unit on either the orthographic or the phonological dimension. Morphological structure in Hebrew is thus fairly opaque. Yet, we demonstrate that Hebrew speakers are sensitive to a phonological constraint that specifically concerns the structure of the root morpheme. The domain to which this constraint applies, the set of all possible Hebrew roots, is very large, and cannot be defined by any orthographic, phonological or semantic features. Despite the fact that Hebrew roots do not share any distinctive nonmorphemic features, subjects nevertheless treat them as a single linguistic class that is subject to a common constraint. This finding, we believe, supports the view of the root morpheme as a discrete constituent in the representation of Hebrew words.

Before describing our evidence, a brief exposition of some of the central properties of Hebrew morphology is in order. Two of the properties of Hebrew morphology, its productivity and nonlinearity, are potentially important for the representation of morphological constituency. We then review existing empirical evidence for morphological decomposition in Hebrew. Finally, we describe the Obligatory Contour Principle as a means for investigating the structure of morphologically complex words in Hebrew.

1.1. The productivity and nonlinearity of Hebrew morphology

Like other Semitic languages, Hebrew morphology is both highly productive and

Table 1
An illustration of some of the derivations of the root [ktb]

Phonological form	Meaning	Word pattern
katav	wrote	KaTaB
hitkatev	corresponded with	hitKaTeB
ktiv	spelling	KTiB
mixtav	letter	miKTaB

Note. The roots on the left are generated by inserting the coot pattern [ktb] in the word patterns on the right. These word patterns specify the location of each of the root's three consonants and the vowels.

nonlinear. Hebrew words are formed by inserting a root morpheme, an abstract sequence of generally 3 consonants, in a word pattern containing vowels, and sometimes, additional consonants. The root itself is not an independent word, but it may be realized in several words that are morphologically (and often semantically) related. These words are generated by inserting the root into one of several verbal and nominal word patterns, called binyanim and mishkalim, respectively. For instance, the root [ktb] conveys the general meaning of writing. This root may form the verb *katav, he wrote*, by inserting it into a prosodic pattern that is characteristic of the third person, masculine, singular, past tense form in binyan qal (i.e., $C_1aC_2aC_3$; see Table 1). Similarly, the noun *mixtav*, letter, is formed by inserting the root [ktb] in a $miC_1C_2aC_3$ noun template. Each of these words, in turn, may further be subjected to a rich inflectional system. Specifically, the root [ktb] may be conjugated in four of the seven verb patterns, binyanim, and each such binyan could yield approximately 30 inflectional forms (Aronoff, 1994, p. 124). Thus, the conjugation of the verb [ktb] alone yields about 120 distinct, morphologically related, words[1].

This brief example is sufficient to illustrate the productivity of Hebrew morphology. In view of this rich productivity, a representation of Hebrew words in reference to their common root would appear to save considerable rote learning and storage space. However, a second characteristic of Hebrew morphology, its nonlinearity, may prove an obstacle for morphological decomposition. In Hebrew, the root morpheme and the word pattern are not linearly discrete units. Instead, they are interwoven, temporally co-occurring entities. Thus, root consonants are often interrupted by a series of vowels, and sometimes, additional nonroot consonant, provided by the prosodic template. The structure of Hebrew words is well captured by nonlinear autosegmental theories of phonology. Autosegmental theories of phonology represent phonological constituents on distinct levels of representation, i.e., planes. These planes are interconnected by the skeleton, a sequence of timing units. In our example, the root [ktb] is represented on a single plane, whereas the vowels are represented on a separate plane. These planes are interconnected by the skeleton, which specifies the word patterns of *katav* and *mixtau* (see Fig. 1).

[1]The status of the mishkalim, noun patterns, as default inflectional classes is less certain (see Aronoff, 1994; Plunkett and Nakisa, 1996).
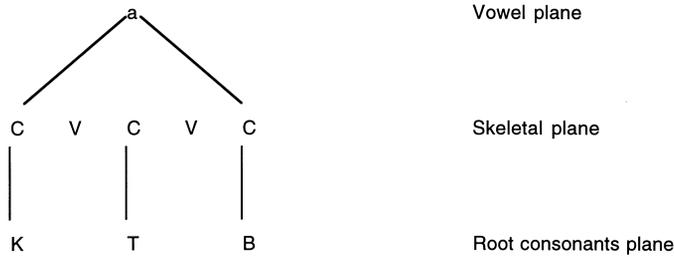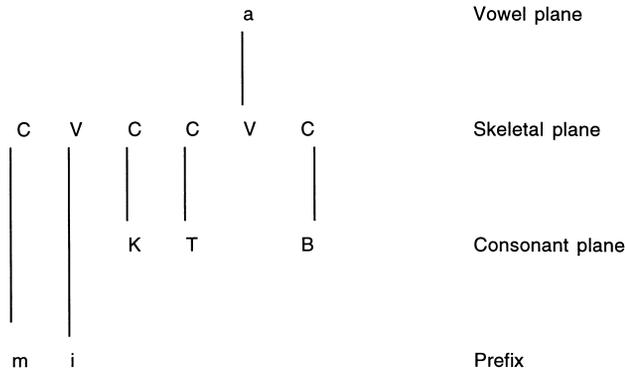
a. Katav

b. Mixtav

Fig. 1. The representations of the words *katav* and *mixtav*. Each representation specifies a skeleton, a root-consonants plane and a vowel plane. Root consonants are thus represented as a single constituent separately from the vowels and any non-root consonants.

This representation neatly captures the fact that, despite their temporal discontinuity, root consonants and vowels each form distinct morphemic constituents. At the same time, however, it illustrates the potential problems posed for the processing of morphological structure in Hebrew. Because Hebrew morphemic constituents often do not correspond to linearly discrete units in either spoken or written language, the root morpheme cannot be easily parsed by simple cues of transitional probability, often available in concatenated morphologies (Seidenberg, 1987). The transparency

of the root is further reduced by phonological processes that obscure the similarity between the root and its derivations. For instance, the words *mixtav* (letter), and *katav* (wrote), are both derived from the root [ktb], whose initial radical is realized as /x/ as a result of a spirantization rule (Glinert, 1989)[2]. Thus, not only does the root morpheme in Hebrew lack linear discreteness but it is also fairly opaque phonologically. How, then, do Hebrew readers represent morphologically complex words? In particular, does the representation of Hebrew words decompose the root and the word pattern?

1.2. Evidence for morphological decomposition: The processing availability rationale

Despite the opaqueness of morphological constituents in Hebrew, there is considerable empirical evidence suggesting morphological compositionality in Hebrew (for a review, see Bentin and Frost, 1992). Several recent studies specifically address the role of the root as a morphemic constituent among adult Hebrew speakers (Bentin and Feldman, 1990; Feldman and Bentin, 1994; Feldman et al., 1995; Frost et al., 1994). Underlying these works is a common rationale concerning the *availability of the root to mental process*. If the root consists of a distinct constituent in mental representations, then the root should be more readily available for processing compared to a sequence of consonants that is not a morphemic constituent. The processing availability of morphological constituents should further increase with the productivity of the root: Highly productive roots should be more easily decomposed from their word patterns compared to unproductive roots. In a segment shifting task, Feldman et al. (1995) instructed subjects to strip the word patterns from a Hebrew word and shift it to a novel root. Feldman et al. (1995) found that the productivity of the root facilitated performance, regardless of its orthographic transparency and surface frequency. This finding suggests that productivity increases the availability of both the root and the word pattern to mental processes.

A closely related argument for compositionality is based on the contribution of root priming to word processing. If the root is a mental constituent, then increasing its availability by priming should contribute to word processing more than a non-constituent sequence of consonants. Supporting this view are the results of several priming studies demonstrating that the priming of the target by a prime sharing the target's root facilitates processing. This facilitation was obtained regardless of the semantic relatedness between the prime and the target, and even when the target and the prime were separated by 15 items (Bentin and Feldman, 1990). In a subsequent study, Feldman and Bentin (1994) observed morphological priming for derivationally and inflectionally related targets and primes, similar in its magnitude to identity priming. Supporting the automaticity of these priming effects, a facilitation by word prime sharing the target's root was obtained regardless of the type of foils (Feldman and Bentin, 1994) and even when the prime was heavily masked (Frost et al., 1994).

---

[2] A similar process accounts for the surfacing of the third root radical, /b/, as /v/ in both *katav* and *mixtav*.

These studies suggest that the root is more readily available to mental processes than non-root controls. A permanent or temporary increase in the availability of the root (due to high productivity or priming, respectively) facilitates word recognition. This finding is readily explained by assuming morphological decomposition. However, other explanations for the increased availability of the root are possible as well. Hebrew roots are strongly associated with a core meaning that is often shared with the target word. Furthermore, because the Hebrew orthography is consonantal, the root morpheme is often represented by a linearly discrete orthographic unit. Thus, in order to attribute the increased availability of the root to its morphological status, it is necessary to rule out the possibility that its greater availability is in fact due to orthographic, phonological and semantic factors. Although existing studies have invested considerable effort in ruling out such nonmorphemic factors, their empirical control cannot be easily achieved. The elimination of semantic, orthographic and phonological explanations for morphological effects is typically inferred from the manipulation of these variables in separate studies, rather than from directly controlling all variables within a single design (Stolz and Feldman, 1995). Evidence for the separate morphemic status of the root could thus be strengthened by the convergence of findings from distinct theoretical perspectives and empirical manipulations.

1.3. The Obligatory Contour Principle: Converging evidence for morphological decomposition

In the present study, we provide converging evidence for the representation of the root as a separate morpheme. The rationale guiding our demonstration is different than that of existing research. Rather than probing for the effect of morphological constituency on the ease of mental processing, we assess Hebrew speakers' tacit knowledge regarding the structure of the root morpheme. We exploit the fact that Hebrew roots are subject to a phonological constraint that is specifically sensitive to location within the root morpheme. Importantly, this constraint cannot be explained by merely referring to the location of the segment in the word or syllable. We consider this formal constraint on the structure of the root morpheme as evidence for morphological decomposition. If Hebrew speakers' knowledge of their native language entails a constraint that specifically concerns the root, then, in order to account for their competence, it is necessary to postulate that the root is a separate morphemic constituent.

The constraint in question concerns the gemination of root consonants. Semitic languages allow a gemination of adjacent root consonants, but constrain its location: A gemination of the second and third root consonants (e.g., [smm]) is frequent whereas a gemination of the first and second root consonants (e.g., [ssm]) is rare (Greenberg, 1950). In his seminal work, McCarthy (1979) provided an elegant explanation for this asymmetry within the framework of autosegmental phonology. McCarthy (1979) departed from the proposal that Semitic languages represent the root consonants on a separate autosegmental plane. He further assumed that the root plane is constrained by the Obligatory Contour Principle (OCP), a

a   a   vowel plane

C V C V C ⟶ C V C V C   skeleton

s   m   s   m   root consonants
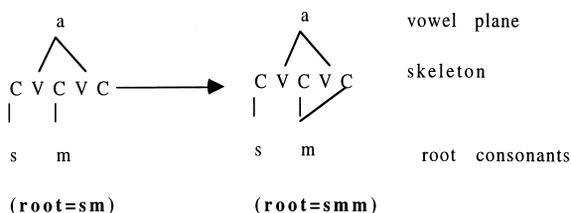
(root=sm)   (root=smm)

Fig. 2. The structure of the verb **samam** and its derivation. The left figure illustrates the underlying representation of the verb, containing a biconsonantal root, **sm**, a skeleton, specifying three consonant positions, and the vowel melody. The right figure demonstrates its surface representation, whose triconsonantal root, **smm**, contains a geminate. This geminate is due to the spreading of the radical *m* to the adjacent empty consonant position in the skeleton.

principle initially proposed by Leben (1973) in studying tonal languages and subsequently documented in a wide variety of phenomena (e.g., Goldsmith, 1990; Kenstowicz, 1994; McCarthy, 1986; Yip, 1988; Yip, 1989). The OCP prohibits the occurrence of adjacent identical elements on the same plane. A gemination of a root consonant is thus banned from underlying representations. The ubiquitous, root-final gemination, (e.g., [smm]), as well as the rare, root-initial gemination (e.g., [ssm]) are considered surface manifestations of a common underlying biconsonantal representation (e.g., [sm]). To derive a word, the root must be associated with a skeletal template that is characteristic of a particular verb or noun pattern, binayan or mishkal. Fig. 2 illustrates this process for the verb *samam*. First, the underlying representation of the root, [sm], is associated with the consonant skeletal positions. This association is assumed to proceed from left to right[3]. Hence, the two initial consonant positions are filled, leaving the third one empty. This empty slot is later filled by spreading of an adjacent consonant, resulting in a surface geminate. Importantly, please note that only the second consonant, /m/, is free to spread. The initial consonant /s/ cannot spread because its adjacent slot is already filled. A gemination of the first and second root consonants is indeed rare. In contrast, because the slot adjacent to the second consonant, /m/, is empty, it is free to spread, resulting in the frequent gemination of the second and third radicals, [smm].

Despite the pivotal role of the OCP in explaining a wide variety of linguistic phenomena, the psychological reality of this principle has not been corroborated yet in experimental settings. Thus, although the OCP is observationally adequate in accounting for the statistical distribution of different root types, it is unclear whether this principle actually reflects the linguistic competence of modern speakers of Semitic. Specifically, in the case of modern Hebrew, the psychological reality of the OCP is questioned by the existence of several violations of this principle. Even-Shoshan's New Hebrew dictionary (Even-Shoshan, 1993) lists 12 roots manifesting root-initial gemination, 3 of them are frequently used in modern Hebrew. Furthermore, a sensitivity to co-occurrence of root consonants presupposes the representa-

---

[3]The direction of association refers to the phonemic transcription (in which the initial phoneme leftmost), rather than to the Hebrew orthography (which proceeds from right to left). Thus, a left to right association proceeds from the beginning of the word to its end.

tion of the root as a separate morpheme: Note that the asymmetry predicted by the OCP constrains the adjacency of elements as a function of their morphological structure, not merely their temporal or spatial adjacency. As previously noted, however, the nonlinearity of Hebrew morphology may obscure morphological constituency. Thus, the opaque nature of the root morpheme and the presence of counter-examples may prevent modern Hebrew speakers from internalizing this phonological constraint on root structure.

The goal of the present research is to examine the sensitivity of Hebrew speakers to the constraint on root structure predicted by the Obligatory Contour Principle. We view the psychological reality of the OCP not only as an intriguing research question in its own right, but further, as a means for uncovering the morphological structure of Hebrew words. If it can be shown that Hebrew speakers possess a knowledge that constrains the position of geminates relative to the root, then it follows that the root is a separate constituent in the representation of Hebrew words. Specifically, a constraint commonly affecting all Hebrew roots would suggest that they all form a single linguistic class. Note, however, that there are no semantic, orthographic, or phonological characteristics that define Hebrew roots as a class. The feature that unites all Hebrew roots must then be formal in nature, i.e., their status as a distinct morphological unit. The sensitivity of Hebrew speakers to a restriction whose domain is specifically the root morpheme would thus provide strong evidence for morphological decomposition.

1.4. Assessing subjects' sensitivity to the OCP via the rating task

Our method of choice in examining readers' sensitivity to the OCP is a rating task. The use of this technique in assessing tacit linguistic knowledge is sometimes criticized on the grounds that the rating decision entails a conscious, problem solving component, its outcomes may reflect meta-linguistic knowledge, and its measurement scale is coarse. These arguments were countered by Prasada and Pinker (1993), who noted that although the rating decision is a conscious act, the computations leading to the decision are not. Indeed, Kim et al. (1991) demonstrated that the ratings assigned to past tense inflections reflect sensitivity to grammatical categories of which subjects are utterly unaware. We return to the demonstration of the tacit nature of the OCP in the discussion of Experiment 1. Like Prasada and Pinker (1993), we further believe that in exploring a previously uninvestigated phenomenon such as the OCP, establishing what subjects consider to be ''an acceptable form'' is logically prior to investigating the effect of well formedness on performance in speeded response tasks. Rating appears to provide a direct and simple reflection of subjects' notion of well formedness. Conversely, the influence of task-specific strategies in speeded response tasks is well documented (e.g., Balota and Chumbley, 1984; Stone and Van Orden, 1993), and their sensitivity does not necessarily exceed the rating technique due to fluctuations in attentional and sensorimotor factors.

To assess subjects' sensitivity to the OCP, we presented them with a series of nonwords created from nonroots (combinations of three consonants that do not

Table 2
An illustration of the materials used in Experiments 1–2[a]

Root type	Class 1	Class 2	Class 3
[ssm]	Si-SeM	maS-Si-Mim	hiS-ta-SaM-tem
[mss]	Mi-SeS	maM-Si-Sim	hit-Ma-SaS-tem
[psm]	Pi-SeM	maP-Si-Mim	hit-Pa-SaM-tem

[a] To clarify the items' morphological structure, root consonants are notated in upper case. The hyphens indicate syllable boundaries. Neither of these notational distinctions were present in the experimental materials.

correspond to any existing Hebrew root) and asked them to determine the extent to which they sound like possible Hebrew words. The structure of our materials is illustrated in Table 2. The critical items violated the OCP due to the gemination of root-initial consonants[4]. If subjects are sensitive to the OCP constraint, then non-words created from such roots should receive low acceptability ratings. To secure the attribution of the low acceptability of these items to a constraint on root structure, rather than a general unacceptability of gemination or a particular ill-formedness of the non-geminating consonants, each of the roots with initial gemination was matched to two control roots.

The first control was designed to examine whether the low acceptability of [ssm]-type roots is indeed due to the location of the gemination. Although Hebrew contains numerous roots manifesting final-gemination, it is possible that subjects have a general bias against gemination that is unrelated to its location. To examine this possibility, each of the initial-gemination roots was matched against a control root geminating the second and third consonants (e.g. [mss][5]). This control root had exactly the same phonemes as the critical root and altered only the location of the gemination. A second alternative explanation to the ill-formedness of the [ssm]-type roots may be unrelated to gemination per sé. On this view, this group of roots exhibits some systematic ill-formedness in the sequence of the second and third consonants, and it is this ill-formedness, rather than the initial-gemination, that accounts for their low acceptability. Although the likelihood of creating such a

[4]Our reference to root-initial gemination as a violation of the OCP constraint is an oversimplification. So is our reference to the sensitivity to the location of gemination as a sensitivity to the OCP. In effect, root-initial gemination may emerge *either* from a violation of the OCP (i.e., from representing the tri-consonantal root in the lexicon) or from a violation of the left-to-right association convention. Similarly, sensitivity to root structure presupposes both the OCP as well as the above association convention. A distinction between violations of the OCP and the association convention are important, but falls beyond the scope of this research. Our current goal is to examine the sensitivity to root structure, rather than to establish its precise linguistic explanation. Thus, throughout the paper, our reference to OCP-effects should be interpreted as effects of asymmetry in root structure that are predicted, in part, by the OCP.

[5]The examples illustrating initial vs. final-gemination roots ([ssm] vs.[mss], or [smm]) were not used in the experiment. We chose these examples to assist readers who are familiar with the discussions of the OCP in the linguistic literature, which frequently use these roots as exemplars. In effect, however, both [smm] and [mss], (but not [ssm]) are existing roots in modern Hebrew. None of the roots used to form our materials is an existing root.

systematic bias in a large group of items is rather low, we nevertheless decided to examine this possibility by designing a second control for the critical root. This control maintained the sequence of the second and third consonants of the initial-gemination root, but differed with respect to the first consonant, which was not a geminate of the second (e.g., [psm]). Together, the triplet of the initial-gemination (e.g., [ssm]), the final-gemination (e.g. [mss]) and the no-gemination roots (e.g. [psm]) permits determination of whether Hebrew speakers have internalized a constraint on the location of gemination.

Although a low acceptability of [ssm]-type roots would suggest that root-initial gemination is sufficient for low acceptability, it still leaves open the critical question of whether root-initial gemination is necessary for low ratings. Is it the location of the gemination within the root that accounts for its unacceptability, or could it be explained instead by a general prohibition of geminating the word's initial consonants? To demonstrate subjects' sensitivity to the OCP, and hence, to abstract root structure, it is necessary to show that it is the location of the gemination in the root, rather than in the word, that accounts for the unacceptability of the critical items. Although there is no reason to believe that a word-initial gemination is ill formed, it is nevertheless possible that subjects' sensitivity to the OCP may be determined by the transparency of the root which, in turn, may be affected by its location in the word. The transparency of the gemination may further depend on the particular phonotactics of the word pattern.

To examine the effect of the transparency of the root on subjects' ratings, each root triplet was conjugated in one of three classes of word patterns[6]. In the first class, the roots were presented with no additional prefixes or suffixes (e.g., *SaSaM*[7]). The location of the gemination in this class was therefore highly transparent. However, because, in this class, root-initial gemination is also word-initial, it is impossible to determine from this class alone whether the unacceptability of root-initial geminates reflects the location of gemination in the word or in the root. The distinction between word and root location was achieved in the second and third classes. Both classes presented the root sandwiched in between a prefix and a suffix. A sensitivity to the OCP in such circumstances requires a representation of the root. The second and third classes nevertheless differed in the transparency of the gemination due to their particular phonotactic properties. In the second class, the first and second root consonants were adjacent in the word's surface form (e.g. *maS-Si-Mim*). The failure to separate the geminates by a full vowel appears to increase the ill-formedness of the gemination. In contrast, geminates in the third class were separated by at least a full vowel and sandwiched between affixes (e.g. *hiS-ta-SaM-ti*). Furthermore, a metathesis rule affecting the formation of some of the word patterns in the third class also disrupted the succession of root consonants (Glinert, 1989). In our example, *hiStaSaMti*, the root-initial geminates are separated by the affix /t/. Such items

---

[6]The term "word class" has no technical linguistic significance. It is used here to group word patterns according to a feature that is of specific interest to the present research, i.e., the transparency of morphological structure.

[7]To illustrate the word's structure, root-consonants are notated in upper case. No such visual distinctions were present in the materials presented to the subjects.

permit examination of subjects' sensitivity to the OCP under extreme conditions of opacity.

In summary, the following studies assess subjects' sensitivity to root structure by obtaining ratings for words derived from a set of three matched roots (Please note that subjects were presented only with the derived verbs and nouns, and never with the roots themselves). These roots were each conjugated in three different classes of word patterns which differed in the transparency of the root. Two questions are at the center of this investigation: (1) Do subjects consider derivations of roots with initial-gemination as unacceptable? (2) Does the sensitivity to root structure require a transparent morphological structure? A demonstration of subjects' sensitivity to root structure would support the psychological reality of OCP and the morphological decomposition of Hebrew words.

EXPERIMENT 1

Experiment 1 was designed to examine subjects' sensitivity to the OCP and determine the effect of root transparency on its magnitude. Subjects were presented with triplets of nonwords and asked to determine the extent to which each member of the triplet sounds like a possible Hebrew word. The members of the triplets were all conjugated in one of the three word classes described above. They shared precisely the same derivational and inflectional patterns, and differed only with respect to their root structure. Specifically, these roots exhibited either an initial-gemination, final-gemination or no-gemination. If Hebrew speakers represent the root separately, and if roots are further subject to the OCP constraint, then root-initial gemination is expected to result in low acceptability ratings regardless of the word pattern in question. Thus, we predict that root-initial gemination should receive lower acceptability ratings compared to both root-final gemination and no-gemination. However, subjects' sensitivity to the location of gemination may depend on the transparency of the root and the idiosyncratic phonotactics of the word pattern. Thus, the acceptability of root-initial gemination may be modulated by word class.

2. Method

2.1. Subjects

The rating questionnaire was administered as part of a course at the School of Education of Haifa University. 18 native Hebrew speakers served as subjects. They were all students in the School of Education of Haifa University and received no compensation for their participation.

2.2. Materials

The items submitted to subjects' rating were all nonwords. Although we refer to these items as 'words', it should be realized that neither the roots nor the conjugated

items corresponded to existing Hebrew words. These words were derived by conjugating a set of roots in one of several noun or verb patterns. Thus, the structure of each word should be analyzed in terms of two morphological components: its root structure and word pattern.

2.2.1. Root triples structure

The rating materials were derived from 24 root triplets. These roots were all novel consonant combinations that did not correspond to any existing root in the Hebrew language. The structure of the root triplet was designed to examine the constraint on the location of gemination in the root. Each root triplet contained three members. The first member exhibited a gemination of the root-initial consonant (e.g., [ssm]). This root is expected to yield low acceptability ratings due to the violation of the OCP. However, low acceptability of such items may result also from additional, unrelated reasons. The second and third members of the triples were designed as controls against such alternative explanations. One alternative explanation to the ill formedness of root-initial gemination is a general unacceptability of gemination. To examine this possibility, we used final-gemination roots (e.g., [mss]) as controls. These roots were formed by switching the location of the geminates and the nongeminate consonant in the critical, initial gemination root. Thus, the only difference between the initial and final-gemination roots concerns the location of the gemination. A second alternative explanation attributes the ill-formedness of [ssm]-type roots to the combination of the nongeminating consonants. This possibility was assessed using a no-gemination root [e.g., psm]) as a control. This root contained the same second and third consonants as in the critical, initial-gemination member, but differed in the first consonant. Due to the special behaviour of guttural consonants in Hebrew (Glinert, 1989) we avoided the inclusion of gutturals among any of the root triples.

2.2.2. Word construction

After the construction of the root triples, each member of such a triplet was conjugated in three classes of word patterns. These word classes were chosen to assess the generality of the Obligatory Contour Principle with respect to word patterns. Of particular interest is the sensitivity of the OCP to the transparency of the root. Recall that Hebrew word patterns may insert in the root not only vowels, but also consonants. The existence of such additional consonants may reduce the transparency of the root, and thus, affect subjects' sensitivity to root structure. The first class included verb patterns in qal, pi'el and pu'al in the third person, masculine, singular, past form. These words were derived by inserting only vowels in the roots. Thus, the root consonants were highly transparent. In the second and third classes of words additional nonroot consonants were affixed to the root. Thus, the second and third classes reduced the surface transparency of the root. In both groups, the initial consonant of the word was a prefix, rather than part of the root. They differed, however, in their phonotactic properties. The word patterns in the second class were characterized by the absence of a full vowel between the first and second root consonants. These word classes included the mishkalim maf'il,

mif'al, hif'il, nif'al and taf'il. Words in the third class were all generated in the hitpa'el past tense form. Thus, the second and third root consonants were always separated by the full vowel *a*. The presence of a root-initial sibilant in the hitpa'el, the binyan used to derive our third class words, triggers a metathesis rule which switches the prefix consonant /t/ with the root-initial consonant (e.g. *hiStaSaMti*, see Glinert, 1989). Thus, the transparency of words formed from roots with an initial sibilant is further reduced in the third class due to a disruption in the order of the root consonants.

In summary, then, three classes of word patterns which differed in the transparency of the root consonants were used in this experiment. Each member of the root triplet was conjugated in each of these three classes. All members of a given triplet were conjugated in precisely the same mishkal or binyan and the same tense and part of speech. Thus, all members of a triplet in a given word class had precisely the same non-root consonants and vowels, and they differed only in their root consonants. The conjugation of each member of a root triplet in each of the three word classes resulted in 9 related word forms for each item. The conjugation of 24 roots in each of the 3 classes of word patterns resulted in 72 word triplets, and a total of 216 words.

2.3. Procedure

Subjects were presented with a printed list of 72 word triplets. The orthographic representation of these words specified all their vowels using diacritic marks. Subjects were asked to rate the extent each member of the triplet sounded like a possible Hebrew words. The order of the presentation of the 72 word triplets was randomly determined. Similarly, within any given triplet, the three members of the triplet were arranged in a random order. Subjects were run in a group. The instructions for the subjects were as follows: 'In this experiment, we attempted to invent new Hebrew words and to find out which of these words sound the best. The words on the following pages are arranged in triplets. Your task is to read each word triplet and to rate within it which word sounds most possible, which one sounds less possible, and which one sounds the least possible. You are asked to read each triplet a couple of times and to silently pronounce the word as it is written. Then, if the word sounds the best, write near it the number 1, if the word sounds the worst, write near it the number 3, and if it is between the other two, write near it the number 2. None of these words is an existing Hebrew word. Do not attend to the meaning these words may remind you. Address only their *sound* and ignore any other aspect. Within each triplet, write 1 near the word that sounds the best, 2 near the word that sounds intermediate, and 3 near the word that sounds the worst'. To express high acceptability by larger numbers, we report the data using an inverted scale, created by subtracting each score from 4. Thus, in our report, 1 corresponds to the worst sounding items whereas 3 to the best.

2.3.1. Design

The root structure (3) and word class type (3) were within subjects and within items variables.

3. Results

Subjects' ratings were submitted to ANOVA's (3 root type × 3 word class) by subjects and items. The ANOVA indicated a significant main effect of root type ($F_1(2, 34) = 111.22$, $p = 0.0001$; $F_2(2, 46) = 174.82$, $p = 0.0001$) and a significant interaction of root type × word class ($F_1(4, 68) = 23.16$, $p = 0.0001$; $F_2(4, 92) = 10.17$, $p = 0.0001$). The main effect of word class was not significant ($F_1(2, 34) = 1.43$, $p = 0.25$, n.s.; $F_2(2, 46) = 1.47$, $p = 0.24$, n.s.). The effect of root type was further investigated using two planned orthogonal contrasts comparing the acceptability of root initial gemination relative to final-gemination and no-gemination controls. Subjects' mean ratings as a function of root type and word class are plotted in Fig. 3.

Across word classes, initial-gemination [ssm] roots were rated significantly lower compared to either final-gemination ($\Delta = 0.682$, $t_1(34) = 8.97$, $p = 0.0001$; $t_2(46) = 11.28$, $p = 0.0001$) or the no-gemination roots ($\Delta = 1.122$, $t_1(34) = 14.79$, $p = 0.0001$; $t_2(46) = 18.55$, $p = 0.0001$). Thus, native Hebrew speakers are sensitive to the location of the gemination in the root. Tukey HSD comparisons performed separately in each of the word classes further confirmed the generality of this effect. Derivations of [ssm]-type roots were rated significantly lower compared to derivations of [mss]-type roots in the first ($\Delta = 0.525$); the second ($\Delta = 1.162$); and the third ($\Delta = 0.359$); class in both the subjects (all $p < 0.01$) and item (all $p < 0.05$) analyses. Similarly, derivations of [ssm]-type roots were less acceptable than derivation of [psm]-like roots in the first ($0 = 0.944$), second ($\Delta = 1.484$) and third ($\Delta = 0.937$) class by subjects (all $p < 0.01$) and items (all $p < 0.0.5$).
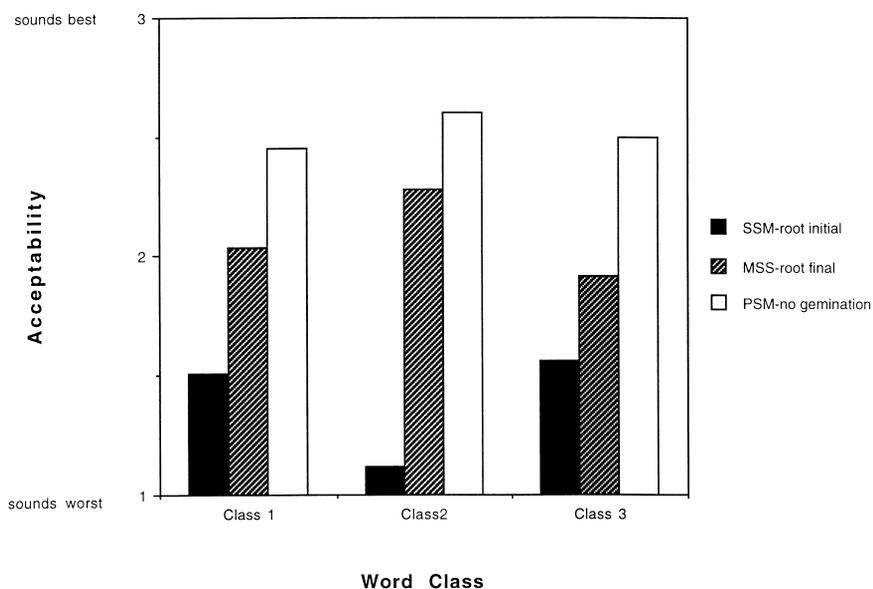

Fig. 3. Mean ratings in Experiment 1 as a function of root type and word class.

Although a sensitivity to OCP violations emerged in each of the word classes, its magnitude was modulated by the word pattern, resulting in a significant interaction of root × word class. A post hoc investigation of this effect was carried out using unplanned contrasts. The unacceptability of [ssm]-type roots compared to [mss]-type roots was significantly greater in word patterns in the second class compared to either the first ($\Delta = 0.637$, $F_1(8, 68) = 5.967$, $p = 0.0000$; $F_2(8, 92) = 2.623$, $p = 0.0124$) or the third class ($\Delta = 0.803$, $F_1(8, 68) = 9.504$, $p = 0.000$; $F_2(8, 92) = 4.175$, $p = 0.0003$). Similarly, the disadvantage of [ssm]-type roots compared to their no-gemination controls was marginally stronger in the second class compared to the first ($\Delta = 0.539$, $F_1(8, 68) = 4.286$, $p = 0.0003$; $F_2(8, 92) = 1.881$, $p = 0.0724$) and the third ($\Delta = 0.546$, $F_1(8, 68) = 4.395$, $p = 0.0003$; $F_2(8, 92) = 1.93$, $p = 0.0646$) classes. Thus, violations of the OCP were especially salient in the second class.

An inspection of the data further suggests a general bias against gemination. This bias was investigated using the Tukey HSD test. Across word classes, the derivation of final-gemination roots received lower ratings compared to no-gemination ($\Delta = 0.44$, $p < 0.05$, by subjects and items). The same bias emerged in each of the three word classes in both the item and subject analyses ($p < 0.05$). Despite the bias against gemination, the ratings of root-initial gemination were nevertheless lower than root-final gemination. Thus, it is the location of the gemination in the root, rather than gemination per se, which accounts for the rejection of the [ssm] type roots.

4. Discussion

The results of Experiment 1 demonstrate that subjects' ratings of nonwords generated from nonexisting Hebrew roots are sensitive to the structure of the root. Words formed by the conjugation of roots with initial-gemination (e.g. [ssm]) were unacceptable. Their unacceptability cannot be attributed to a general bias against gemination, since the ratings of root-initial gemination were significantly lower compared to final-gemination roots (e.g., [mss]). Similarly, the unacceptability of words formed from [ssm]-type roots cannot be attributed to some idiosyncratic ill-formedness residing in the second and third root radicals, as these words were rated significantly lower than controls generated from roots containing these consonants (e.g., [psm]). Thus, the unacceptability of words derived from [ssm]-type roots reflects a sensitivity to the location of gemination. Importantly, subjects' ratings refer to the location of gemination within a *morphological* domain, the root.

The sensitivity to the location of gemination cannot be attributed to its position within the word: Subjects rejected derivations of [ssm]-type roots regardless of the position of the root in the word. Furthermore, the sensitivity to the position of the gemination within the root was general, and its strength was not reduced as a function of the transparency of the root within the word. Derivations of [ssm]-type roots were rejected in the first class (e.g., *SaSaM*), in which the root was highly

transparent and the gemination was word initial; the second class (e.g., *maSSiMim*) in which the root was sandwiched by a prefix and a suffix; and in the third class (e.g., *hiStaSaMtem*), in which the root is heavily affixed and the geminates were separated by a full vowel. An especially strong test of subjects' sensitivity to root structure is presented by third class words derived from roots whose initial radical is a sibilant, such as [ssm]. These roots are subject to a metathesis rule which disrupts the integrity of the root consonants by inserting the affx /t/ (e.g., hiS-ta-SaM-tem, see Glinert, 1989). Our materials included 8 items with a root-initial sibilant which underwent metathesis in the third class. A separate examination of these items reveals the same general pattern obtained in our experiment: Root-initial gemination was rated significantly lower than root-final gemination ($\Delta = 0.688$, $t(7) = 3.12$, $p = 0.017$) or no-gemination ($\Delta = 0.937$, $t(7) = 7.12$, $p = 0.0002$) controls.

   Subjects' sensitivity to root structure is quite striking. To assess its reliability, we attempted to replicate our findings. In a replication of Experiment 1, a new group of 18 subjects rated 198 words derived by conjugating a set of 22 root triplets (12 of which overlapped with those used in Experiment 1)[8] in each of the three word classes described in Table 2. The structure of the materials and the rating procedure were identical to those employed in Experiment 1. The results, described in Fig. 4, essentially replicate Experiment 1's findings. Subjects were highly sensitive to root structure. The ANOVA's (3 type × 3 word class) yielded a significant main effect of root type (F_1 (2, 34) = 71.71, $p = 0.0001$; F_2(2, 42) = 82.37, $p = 0.0001$). Planned comparisons indicated that root-initial gemination was rated lower than either root-final gemination ($\Delta = 0.5312$, $t_1(34) = 6.6554$, $p = 0.0001$; $t_2(42) = 7.1335$, $p = 0.0001$) or no-gemination controls ($\Delta = 0.9537$, $t_1(34) = 11.946$, $p = 0.0001$; $t_2(42) = 12.807$, $p = 0.0001$). As in Experiment 1, the interaction of root type × word class was significant ($F_1(4, 68) = 9.07$, $p = 0.0001$; $F_2(4, 84) = 8.95$, $p = 0.0001$). A post hoc investigation of this interaction using unplanned contrasts suggested that the unacceptability of root initial gemination was especially strong in the second class. The disadvantage of [ssm]-type roots relative to [mss]-type controls was significantly larger in the second class compared to either the first ($\Delta = 0.684$, $F_1(8, 68) = 3.741$, $p = 0.0011$; $F_2(8, 84) = 3.691$, $p = 0.0010$) or the third class ($\Delta = 0.525$, $F_1(8, 68) = 2.204$, $p = 0.0376$; $F_2(8, 84) = 2.174$, $p = 0.0375$). The second class also yielded a numerically, but not significantly, larger disadvantage of [ssm]-type root relative to its no gemination control (Relative to the first class: $\Delta = 0.399$, $F_1(8, 68) = 1.271$, $p = 0.2730$; $F_2(8, 84) = 1.255$, $p = 0.2781$; Relative to the third class: $\Delta = 0.46$, $F_1(8, 68) = 1.687$, $p = 0.1175$; $F_2(8, 84) = 1.665$, $p = 0.1191$). Importantly, however, root-initial gemination was unacceptable in all three word classes: Root-initial gemination was rated significantly lower than no-gemination controls in each of the three classes (Tukey HSD tests, all $p < 0.01$, by subjects and items). The comparison of root initial gemination to final gemination controls revealed a numerical, but statistically nonsignificant, disadvantage in the first class (Tukey HSD tests, $p > 0.05$,). However, the second and third class each

---

[8]The questionnaire presented to the subjects included 24 root triplets. However, two of these triplets had typographic errors, and hence, excluded from all subsequent analysis, resulting in 22 root triplets.
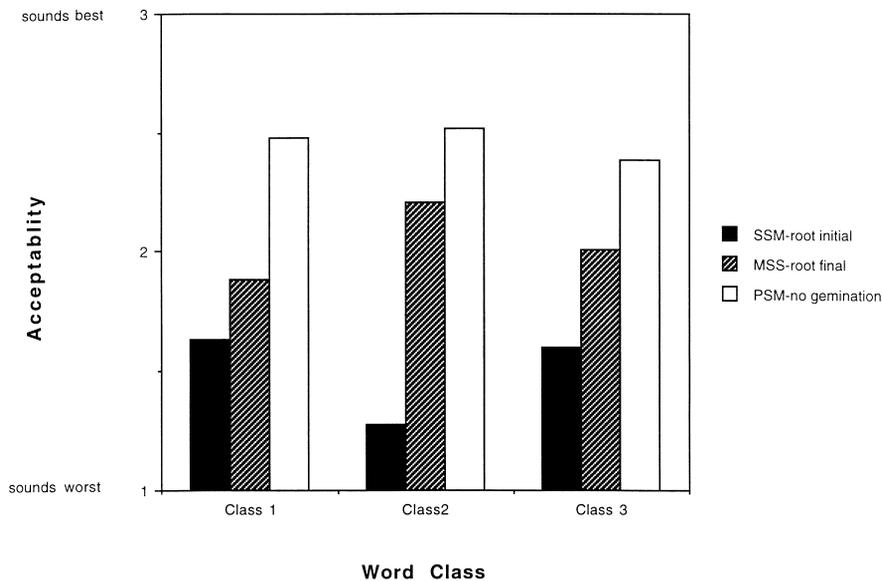
Fig. 4. Mean ratings in the replication of Experiment 1 as a function of root type and word class.

reflected significantly lower ratings for [ssm]-type roots compared to [mss)-type controls (Tukey HSD tests, all $p < 0.01$, by subjects and items). Subjects are thus sensitive to root structure even when the word's morphological structure is opaque, in the second and third class. Additional support for subjects' sensitivity to morphological structure comes from separate analyses conducted on the set of 8 spirants undergoing metathesis in the third class. Recall that, in these items, root structure is especially opaque due to its disruption by an infix (e.g. hiS-ta-SaM-tem). These items nevertheless revealed lower ratings for root initial gemination compared to final gemination ($\Delta = 0.59$, $t(7) = 4.157$, $p = 0.0043$) or no-gemination ($\Delta = 0.722$, $t(7) = 2.923$, $p = 0.0223$) controls.

The sensitivity of subjects' ratings to the location of gemination in the root fits well with the claim that Hebrew roots are decomposed from their word patterns and constrained by a co-occurrence restriction such as the OCP. Before accepting this conclusion, however, alternative explanations must be considered. We noted that rating outcomes are sometimes criticized as unindicative of tacit linguistic knowledge. Specifically, subjects' sensitivity to root structure may not necessarily reflect their morphological representation. The unacceptability of root-initial gemination may be explained away by unfamiliarity with these forms or difficulties in their articulation. Indeed, subjects may use such notions to label their experience. These explanations, however, suffer from circularity. The reason a verb such as *sasamti* (derived from [ssm]) might appear less familiar than *samamti* (derived from [smm]) is not that the specific linguistic token is less frequently encountered. With all likelihood, neither *sasamti* nor *samamti* was ever encountered by our subjects before, nor had they ever encountered their roots. Indeed, all items presented to the subjects were nonwords derived from nonroots. Thus, a perceived unfamiliarity with the

form *sasamti* can only reflect an unfamiliarity with its **structure**. In particular, to explain this percept, it is necessary to appeal to an abstract representation of the root and the location of gemination within it. This is precisely what the OCP principle predicts. Similarly, the unacceptability of *sasamti* cannot be explained away by a simple notion of 'articulatory difficulties'. A consistent difficulty in articulating the surface forms derived from root-initial gemination is unlikely given that their rejection emerged for diverse word patterns and roots. If the data are to be explained by appealing to articulation, then this account must be formulated in reference to an abstract structure, rather than specific surface tokens. Such explanation is not implausible. However, it must minimally presuppose the very same abstract structural description we appeal to, namely, the decomposition of root from word patterns and the location of gemination in the root.

An apparently stronger challenge for our conclusions is one frequently raised regarding rating data, namely, the metalinguistic explanation. Critics of the rating method may agree with us that an account of our findings must presuppose a knowledge which decomposes root structure from word patterns. However, they may disagree with our interpretation of its source as tacit linguistic competence. It is well known that Hebrew children learn at school about the decomposition of roots from their word patterns (Bentin and Frost, 1992). The demonstration of a metalinguistic knowledge regarding root structure obviously may have little bearing on tacit linguistic competence. However, the actual contents of our subjects' metalinguistics knowledge cannot account for the data. Consider again the verbs *sasamti* and *samamti*. The decomposition of these verbs in accord with subjects' metalinguistic knowledge would yield the roots [ssm] and [smm]. In contrast to most existing research in Hebrew morphology, our task does not examine merely the stripping of the root from the word pattern. Instead, we are looking for a behaviour that differentially treats these two roots. A metalinguistic knowledge of decomposition should not differentiate these root patterns. The rejection of [ssm]-type roots must reflect additional knowledge regarding the location of gemination. Importantly, however, we find no evidence that this knowledge is metalinguistic. In contrast to the reflective knowledge of decomposition, which is indeed readily available to Israeli adults, the OCP is a principle known only to trained linguists, and is not taught in the Israeli school system. Similarly, the asymmetry in the distribution of root geminates is not conspicuous. Informal conversations with our subjects indicated they were utterly unable to explain why verbs such as *sasamti* are less acceptable than *samamti*. Thus, subjects appear to lack a reflective knowledge of the OCP or the distribution of root geminates. The fact that their behaviour is nevertheless constrained by knowledge of root structure must indicate that its source is tacit linguistic competence. In particular, to explain the pattern of ratings, it is necessary to assume that subjects' tacit knowledge regarding the location of gemination refers to the root morpheme as its domain. This finding supports the view of the root morpheme as a separate constituent in the representation of Hebrew words. The sensitivity of subjects to the location of gemination within the root is further compatible with the idea that root structure is constrained by phonological co-occurrence restrictions such as those stated by the Obligatory Contour Principle.

Finally, two of our findings require additional explanation. First, subjects exhibited a general bias against gemination. Although root-initial gemination was unacceptable compared to root-final gemination, final-gemination was rated lower than no-gemination. A similar highly significant bias emerged in the replication of Experiment 1. Specifically, across root types, SSM roots were rated lower than PSM type roots ($p < 0.01$, by subjects and items, Tukey HSD). The bias against root final gemination was significant in each of the three word classes (all p's < 0.01 by subjects and items; Tukey HSD tests). This finding is surprising, given that root-final gemination is common in modern Hebrew (e.g., *aa, bdd, brr, gzz, gss, dll, dmm, zll, zmm, xmm, xff, kll, lkk, mss, mrr, ndd, sbb*). A second finding that requires further explanation is the effect of word class on the acceptability of root-initial geminates. The source of these two findings is not entirely clear, but they, may be related to the OCP phenomenon.

McCarthy (1986) demonstrates that the OCP constrains not only lexical representations but also their derivation. Viewed generally, the OCP may reflect an antigemination effect: it blocks phonological processes resulting in tautomorphemic geminates. Antigemination may thus operate at distinct levels of representation and its consequences may differ in nature and strength. In our data, the rejection of root-initial gemination may reflect an early constraint on lexical representations. Conversely, the bias against gemination and increased rejection of root-initial gemination in the second class may both reflect antigemination effects at later stages of the derivation. Specifically, if antigemination processes may inspect derived representations, then the gemination in roots like [smm] may be noted, and perceived as less preferred compared to no-gemination roots. However, a gemination in the derived, tri-consonantal root, (e.g., [mss]) may be obscured by plane conflation (McCarthy, 1986), a process that results in the representation of the root and vowel melody on a single plane. Because root-final geminates are often separated by a full vowel (e.g., *SaMaM*), a late antigemination effect in surface forms should be weak. Indeed root-final gemination is common. In contrast, the process of plane conflation cannot obscure root-initial gemination in the second class. Forms in the second class, like *MaS-SiMim*, do not separate the geminates by a vowel. Root-initial geminates in these forms are thus truly adjacent even after plane conflation. Such surface adjacency appears highly undesirable, and is often avoided by means of vowel epenthesis (McCarthy, 1986). A late antigemination effect may thus accentuate the ill-formedness of root-initial gemination in the second class, resulting in pronounced rejection of [ssm]-type roots.

The nature of the OCP as an antigemination effect awaits further research. At present, however, our conclusions may be summarized as following. Experiment 1 demonstrated that subjects are sensitive to the location of gemination within the root. The asymmetry in the ratings of root-initial vs. root-final gemination is compatible with the idea that subjects have internalized a constraint on root structure such as the OCP. Conversely, the fact that the root serves as a domain for a mental constraint suggests that it forms a separate constituent in the representation of Hebrew words.

EXPERIMENT 2

Experiment 1 demonstrates the sensitivity of Hebrew speakers to the structure of the root. This finding indicates that Hebrew speakers represent the root morpheme as a separate constituent. Furthermore, in making their acceptability judgments, subjects must attend to the root by stripping its affixes. How readily do Hebrew speakers attend to root-structure?

The fact that subjects are highly sensitive to the structure of the root even when it is sandwiched between a prefix and a suffix and its integrity is disrupted by an infix suggests that subjects can strip the root even when word structure is extremely opaque. However, the nature of the rating procedure used in Experiment 1 may somewhat limit the generality of this conclusion. Recall that subjects were instructed to rate each word relative to the other two members in the word triplet. These words were identical in all aspects but their roots. Since the root was the only dimension that permitted discrimination between items, it is possible that subjects' attention to the root was encouraged by the rating procedure. Note, however, that this possibility can only limit our conclusions with respect to the generality of root stripping. It does not undermine the conclusion that Hebrew speakers represent the root as a separate constituent in their long term knowledge. Our results clearly demonstrate that subjects possess a knowledge regarding the structure of Hebrew roots. Obviously, such knowledge could not have been acquired solely from the performance of the experimental task. Nevertheless, the generality of root-stripping is important for revealing the processing of morphologically complex words. This question is investigated in Experiment 2.

Experiment 2 examined subjects' sensitivity to root-structure under conditions that did not encourage attending to the root. In this experiment, subjects were presented with the same materials employed in Experiment 1, and they were instructed to rate the words for acceptability on a scale of 1–5. The order of the words within the list was completely randomized. Thus, there is nothing in the new procedure that directs subjects' attention to the root. In principle, subjects could ignore root structure and base their judgment on purely idiosyncratic properties of an item, such as word class and phonotactics. If subjects' attention to root structure in Experiment 1 was an artifact of the rating procedure, then no effect of root structure should be obtained in the present study. However, if word decomposition into root and affixes is mandatory, and if subjects further possess general knowledge regarding the constraints on root structure, then words derived from roots manifesting initial-gemination should receive low acceptability ratings regardless of word class.

5. Method

The same set of 216 non-words employed in Experiment 1 was used in the present study. These words were formed by the conjugation of the 24 root triplets used in Experiment 1 within each of the 3 word classes described previously. The words

were presented in a randomized list. The instructions for the subject were similar to those described in Experiment 1 with the following exceptions. First, subjects were asked to rate each word according to how it sounded individually, rather than in reference to any other words. Second, the rating scale was changed to a 5 point scale. This scale indicated the extent to which the word sounds like a possible Hebrew word. 5 indicated a word that sounds excellent, 4 indicated a word that sounds good, 3 indicated a word that sounds strange, 2 indicated a word that does not sound good, and 1 indicated a word that sounds impossible.

5.0.1. Subjects

15 native speakers (Haifa University Psychology Students) served as subjects. The rating questionnaire was administered as part of a course. Subjects were not compensated for their participation.

6. Results

Subjects' ratings were submitted to analyses of variance by subjects and items (3 root type × 3 word class). The main effects of root type (F_1 (2, 28) = 63.34, $p = 0.0001$; F_2(2, 46) = 75.03, $p = 0.0001$), word class (F_1(2, 28) = 16.50, $p = 0.0001$; F_2(2, 46) = 39.44, $p = 0.0001$) and the interaction of root type × word class (F_1(4, 56) = 27.31, $p = 0.0001$; F_2(4, 92) = 41.69, $p = 0.0001$) were all significant.

Subjects' mean ratings as a function of root type and word class are presented in Fig. 5. Subjects' sensitivity to the location of gemination was investigated by two orthogonal planned contrasts. In accord with the results of Experiments 1, across word classes, subjects rated words derived from roots with initial-gemination significantly lower compared to either final-gemination ($\Delta = 0.801$, $t_1(28) = 9.249$, $p = 0.0001$; $t_2(46) = 9.984$, $p = 0.0001$) or no-gemination ($\Delta = 0.881$, $t_1(28) = 10.18$, $p = 0.0001$; $t_2(46) = 11.139$, $p = 0.0001$) controls.

A series of Tukey HSD tests were performed to assess the generality of this effect within each of the word classes. The rejection of root-initial gemination was robust even for the second and third classes, in which the root was affixed. Specifically, in the second class, root-initial gemination was rated significantly lower compared to either final-gemination ($\Delta = 1.603$, $p < 0.01$, by subjects and items) or no-gemination ($\Delta = 1.802$, $p < 0.01$, by subjects and items). Similarly, the rejection of root-initial gemination was obtained in the third word class compared to either root-final gemination ($\Delta = 0.548$, $p < 0.01$, by subjects and items) or no-gemination ($\Delta = 0.647$, $p < 0.01$). These findings suggest that the constraint on gemination refers to the root, rather than the word, and that root stripping takes place even when the word's morphological structure is highly opaque. In fact, it is the first class, in which the root was highly conspicuous, that manifested the weakest rejection of initial-gemination. In the first class, the ratings of [ssm]-type roots were numerically lower than either [mss]-type roots ($\Delta = 0.253$) or no-gemination controls ($\Delta = 0.197$), but this trend did not reach significance by subjects or items ($p > 0.05$, n.s.). Indeed, the
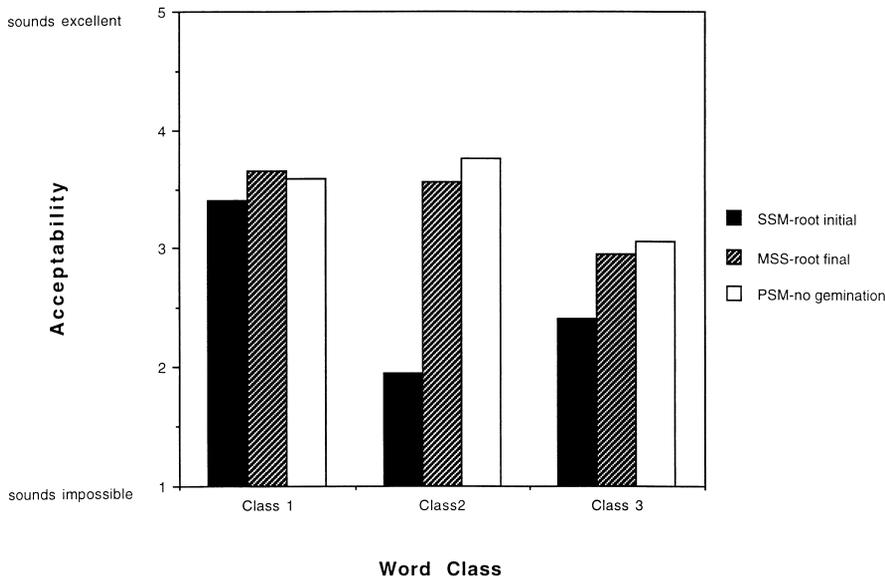
Fig. 5. Mean ratings in Experiment 2 as a function of root type and word class.

interaction of word class × root type indicates a modulation of the effect of root type by word pattern properties. A post hoc investigation of this interaction was carried out using unplanned contrasts. Replicating the results of Experiment 1, the rejection of root-initial gemination in the second word class was especially robust, and its magnitude relative to root-final gemination was significantly greater compared to both the first ($\Delta = 1.35$, $F_1(8, 56) = 7.752$, $p = 0.0001$; $F_2(8, 92) = 11.587$, $p = 0.0001$) and third class ($\Delta = 1.055$, ($F_1(8, 56) = 4.729$, $p = 0.0002$, $F_2(8, 92) = 7.712$, $p = 0.0001$). Similarly, the disadvantage of [ssm]-type roots relative to [psm]-type roots in the second class was larger than in both the first ($\Delta = 1.605$, $F_1(8, 56) = 10.938$, $p = 0.0001$; $F_2(8, 92) = 16.583$, $p = 0.0001$) and third class ($\Delta = 1.155$, ($F_1(8,56) = 5.659$, $p = 0.0002$, $F_2(8, 92) = 8.59$, $p = 0.0001$).

The present results differed from the findings of Experiment 1 in two respects. First, there was no evidence for a general bias against gemination. An evaluation of this bias was carried out using a Tukey HSD. Across word classes, the rating of root-final gemination did not differ from no-gemination ($\Delta = 0.081$, $p > 0.05$, n.s. by subjects and items). Similarly, there was no difference between the ratings of root-final gemination and no-gemination in the first ($\Delta = 0.056$, $p > 0.05$, n.s.), the second ($\Delta = 0.199$, $p > 0.05$, n.s.), or the third class ($\Delta = 0.099$, $p > 0.05$, n.s.). The second new finding in this study is the main effect of word class. Tukey HSD comparisons revealed higher ratings for the first class compared to either the second ($\Delta = 0.459$, $p < 0.05$, by subjects and items), or third class ($\Delta = 0.737$, $p < 0.05$, by subjects and items). The difference in ratings of the second and third class was significant by items only ($\Delta = 0.278$, $p < 0.05$). Thus, subjects generally viewed conjugations in the first class as most acceptable.

7. Discussion

The results of Experiment 2 replicate the central findings of Experiment 1. Despite the fact that the rating procedure did not require special attention to the root, subjects nevertheless manifested a strong sensitivity to its structure. Words derived from roots with initial-gemination were unacceptable compared to roots with final-gemination or no-gemination. The sensitivity to root structure was not contingent on the transparency of morphological structure. In fact, it is the first class, in which the root structure is most conspicuous, that failed to yield a significant effect in this study. The interpretation of this null effect is uncertain. It may be partly due to a lesser sensitivity of the rating procedure used in this study, as a significant effect of root structure did emerge in Experiment 1. Conversely, the instability of the effect in the first class may suggest that subjects are more tolerant of root-structure violations in the first class. This tolerance, however, cannot be due to their inability to strip the root specifically in the first, most transparent class. Indeed, the rejection of root-initial gemination in the most opaque class, the third class, was highly significant and numerically larger than in the first class, in which the root was not affixed. Furthermore, a separate analyses of the 8 roots whose initial radical was a sibilant, and thus, their conjugation in the third class disrupted the sequence of root consonants by an infix, revealed a strong rejection of root-initial gemination compared to either root-final gemination ($\Delta = 0.683$, $t(7) = 2.592$, $p = 0.036$) or no-gemination ($\Delta = 0.942$, $t(7) = 4.588$, $p = 0.002$). This finding reaffirms our previous conclusions that subjects' linguistic competence constrains the structure of Hebrew roots. In addition, it demonstrates that subjects attend to the root even when its structure is extremely opaque, under circumstances in which morphological decomposition is not required or encouraged by the experimental task. These results converge with findings obtained using on-line methods (e.g., Feldman and Bentin, 1994; Frost et al., 1994) suggesting that root stripping may be mandatory.

The new rating procedure reflected some additional aspects governing the acceptability of nonwords. Specifically, this procedure revealed a marked sensitivity to word class. Words formed from the first class were rated significantly higher than in the second or third class. We first consider a linguistic explanation for this pattern. It is possible that the higher ratings of first class verbs reflects a general preference regarding the formation of new Hebrew verbs. Indeed, Bolozky (1982; cited in Aronoff (1994, p. 130)) observed that the distinct binyanim differ with respect to the frequency in which they are used in generating new Hebrew verbs. The pi'el is more productive than the hitpa'el, which, in turn, is more productive than the qal. Recall that in our materials, class 1 was composed of derivations in the pi'el and qal binyanim whereas class 3 consisted entirely of forms in hitpa'el. To examine the correspondence of our ratings to the productivity of the different binyanim, we performed a separate analysis of the verbs in the first and third class according to their respective binyanim. The results of this analysis are presented in Table 3. In agreement with the productivity of the binyanim in Hebrew, our subjects rated verbs formed from binyan pi'el higher (i.e., more acceptable) than verbs in the hitpa'el binyan. In contrast with the productivity of the binyanim, however, the rating of

binyan qal was in fact higher than the hitpa'el. Conversely, Bolozky observes greater productivity of hitpa'el compared to qal in generating new Hebrew words. The correspondence between our data and the productivity of different binyanim in modern Hebrew is thus only partial.

One may thus wish to entertain a nonlinguistic explanation for these findings. In rating nonwords, subjects may be experiencing a conflict between the nonlexical nature of the roots and the ''wordiness'' of the prosodic patterns in which they are conjugated. Words in the third class are more heavily affixed then the other two classes. Furthermore, the hitpa'el has some well defined semantic attributes indicating a reflexive action. The combination of the ''wordy'' aspects of the word pattern, its affixes and meaning, may conflict with the nonlexicality of the root, and appear strange. Conversely, word patterns in the first class, which lack consonant affixes, may appear less ''lexical'', and thus, spared from the conflict with their roots. As a result, first class items may be rated as more acceptable. One may further speculate that the absence of consonant affixation in first class items may reduce their appearance as linguistic objects. Consequently, subjects may be more tolerant of structural violations for such items. This explanation, if correct, could accommodate the null effect of root structure in the first class.

An additional consequence of the new rating procedure employed in this experiment is the elimination of the general bias against gemination observed in Experiments 1. Although this bias was unexpected, it may be explained by an antigemination effect on the derivation of phonological representations. The elimination of this bias with the change in the rating procedure is puzzling. Solving a divergence between the outcomes of experimental methods requires an interpretation of their properties. Unfortunately, such assessment cannot be made in an a-theoretical fashion (for discussion, see Berent and Van Orden, 1996; Van Orden et al., 1996). Indeed, one can easily construct two equally plausible, but mutually exclusive, scenarios that ratify the conclusions of each of the rating methods. On one view, the rating procedure employed in Experiment 1 better directs subjects' attention to root structure. This method may thus be considered more sensitive than the open ended procedure employed in Experiment 2. This interpretation would lead to accept the bias against gemination as a true marker of subjects' linguistic competence. Conversely, one may part from the assumption that the method in Experiment 1 encourages a deliberate problem solving strategy in order to distinguish between otherwise highly similar items. The open-ended procedure in Experiment

Table 3
Mean ratings in Experiment 2 as a function of root type and binyan

Root type	qal	pi'el	pu'al	hitpa'el
[ssm]	3.453	3.438	3.276	2.411
[mss]	3.653	3.981	3.343	2.919
[psm]	3.687	3.676	3.381	3.058
mean	3.598	3.698	3.333	2.796

note: 1 = sounds impossible; 5 = sounds excellent.

2 would be considered free from such artifacts, providing a better reflection of their true linguistic competence.

The existing data cannot discriminate between these scenarios. We are thus unable to resolve this divergence between the two rating methods. More interesting, however, is their convergence: Despite the fact that the rating procedure in Experiment 2 did not direct attention to the root, subjects nevertheless based their ratings on an abstract knowledge of its structure. This conclusion demonstrates that subjects can readily decompose Hebrew words by stripping their roots. Their sensitivity to root structure further indicates that they possess a long term knowledge whose domain is the root morpheme. Thus, the root appears to form a separate constituent in the representation of Hebrew words.

8. General discussion

Hebrew speakers consider words derived from roots with initial-gemination as unacceptable. This finding clearly does not stem from a general bias against gemination[9], as the rating of root-initial gemination was lower than that of root-final gemination. Thus, it is not merely the presence of gemination but its location that accounts for subjects' behaviour. Importantly, the sensitivity to the location of gemination is defined relative to a morphological unit, the root. The rejection of words derived from roots with initial-gemination was obtained regardless of the position of the root in the word, even for words whose morphological structure was extremely opaque due to affixation and root internal infixes.

Subjects' rejection of root-initial gemination has several implications. The first concerns the status of the root as a morphemic unit. Our findings coincide with the conclusions of existing studies (Bentin and Feldman, 1990; Feldman and Bentin, 1994; Feldman et al., 1995; Frost et al., 1994) in providing strong support for the morphological compositionality of Hebrew words. They demonstrate that Hebrew speakers readily attend to the root and strip it from the word pattern. Root stripping takes place even when the root is not an uninterrupted orthographic unit, it carries no meaning, and the decomposition is not directly necessary for the performance of the experimental task. These findings are incompatible with the view of morphological decomposition as an artifact of semantic and orthographic confounds (Seidenberg, 1987; Seidenberg and McClelland, 1989). Please note that, although our evidence is obtained from nonwords, its implications are not limited to the representation of nonwords. We infer morphological decomposition from the behaviour of the root as the domain of a phonological constraint. This constraint, however, can only be explained by assuming that subjects possess a general knowledge regarding the

---

[9]Viewed generally, the OCP may be considered as an antigemination effect (McCarthy, 1986). We do not rule out this view. In fact, it is supported by the unacceptability of root-final gemination in Experiment 1. Instead, we note that a general bias against gemination in surface forms cannot explain the asymmetry in root structure. Thus, a general antigemination effect may apply at different levels of representation. Its consequences at distinct stages may differ in their nature and strength.

co-occurrence of geminates in the root. If all Hebrew roots serve as the domain of this constraint, then any root, regardless of lexicality, must be represented separately from its word pattern.

A second implication of our findings concerns the structure of phonological representations. The fact that the root consonants serve as the domain of phonological processes suggests that they form a phonological constituent, segregated from intermediate vowels. These findings thus demonstrate that phonological representations in Hebrew segregate root consonants and vowels (McCarthy, 1979). The source for this segregation and its generality are unclear at present. On one view, the segregation of consonants and vowels in phonological representations emerges only when they constitute different morphemes (but see McCarthy, 1989). Conversely, a vast body of psychological evidence regarding the perception and production of both spoken and written language, language acquisition and writing disorders indicate that the segregation of consonants and vowels may be far more general (for a review, see Berent and Perfetti, 1995). The present findings are compatible with the suggestion that consonants and vowels may form distinct constituents in phonological representations (Berent and Perfetti, 1995). Future research is required to assess the contingency of consonant-vowel segregation on morphological structure.

A third implication of our findings concerns the knowledge which constrains the structure of Hebrew root-consonants. Our findings demonstrate that Hebrew speakers internalize a phonological constraint on root structure that can be adequately described by the Obligatory Contour Principle. The OCP is one of the most powerful explanatory principles in modern phonology (for reviews, see Goldsmith, 1990; Kenstowicz, 1994). To our knowledge, these results provide the first experimental support for the psychological reality of this principle. At the heart of the OCP, however, is a distinction between the behaviour of geminates and nongeminate bigrams. This proposal has some potentially deep implications for our understanding of cognitive architecture. We review them in the following sections.

8.1. Do mental representations encode gemination?

Our results suggest that Hebrew speakers internalize a constraint regarding the location of gemination within the root morpheme. What is required from a psychological theory to account for these findings? What are the properties of mental representations that are necessary to capture the structure of Hebrew roots?

Symbolic and connectionist models of cognition offer different replies to these questions. We do not intend to evaluate the principled adequacy of these architectures nor do we wish to examine the performance of any specific model. We are not aware of any model designed to handle the OCP phenomenon. More importantly, the specific cognitive hypotheses embodied in a given simulation may not necessarily follow from the properties of its architecture. Specifically we recognize that the theoretical claims embraced by many existing connectionist models are not inherent to connectionism (Marcus et al., 1995). Indeed, we show that a connectionist network can easily capture the critical aspects of our data. The crucial question from the cognitive standpoint is not whether a model can mimic human behaviour, but

instead, what it takes to make the model do so. The principles embodied in a model's design constitute a cognitive hypothesis.

Following Marcus et al. (1995), we consider two strong hypotheses regarding the nature of linguistic knowledge: the symbolic hypothesis vs. the pattern-associator. The symbolic hypothesis assumes that the representation of linguistic knowledge includes mental rules. The crucial aspect of this proposal is not the explicitness of the rules nor the probability of association between their inputs and outputs. Many linguistic rules are implicit; some of them, including the OCP, are soft, violable constraints (Goldsmith, 1990; Kenstowicz, 1994); and the number of exceptions may exceed rule-obeying tokens (Marcus et al., 1995). Importantly, mental rules are structure-sensitive processes whose operation is determined by the combinatorial, constituency structure of mental representations (Fodor and Pylyshyn, 1988). The critical aspect of the symbolic hypothesis is that linguistic representations have constituency structure. Conversely, the pattern-associator hypothesis considers constituency structure obsolete. Linguistic knowledge is explained by the formation of rich patterns of associations over features of linguistic tokens.

The OCP phenomenon presents an interesting test of these hypotheses. The power of the pattern associator in expressing linguistic generalizations resides in its ability to extract the correlational structure of linguistic tokens. These patterns are highly informative. Linguistic rules often mark their outputs by some well defined features. These features may help direct the system and the child's attention to the desired generalization. Consider the widely-studied English past tense verbs (Pinker, 1991; Pinker and Prince, 1988; Plunkett and Marchman, 1993; Prasada and Pinker, 1993; Rumelhart and McClelland, 1986). Regular past tense verbs are marked by a suffix. Thus, these forms may be identified by virtue of their association with a linearly discrete unit and some well defined phonological and semantic features. Although the phonogical and semantic correlates of formal structure may be insufficient to capture speakers' knowledge, attending to phonological similarity of tokens may provide an initial approximation for the desired generalization and a bootstrapping device for its acquisition. By contrast, Hebrew root structure is not flagged by any discrete unit nor is it associated with orthographic, phonological or semantic features. No correlates of formal structure would provide even an approximation of the set of final-gemination roots in Hebrew (e.g., *att, bdd, brr gzz, gss, dll, dmm, zll, zmm, xmm, xff, kll, lkk, rss, mrr ndd, sbb*).

Encoding the constraint on root structure has several prerequisites. One is to represent the root as a separate constituent. It is unclear whether this requirement could be reconciled with the pattern associator hypothesis. Hebrew root structure may not be easily identified by the association of nonmorphemic correlates. Let's assume however, that a pattern associator (i.e., a system whose implementation follows the pattern associator hypothesis) is provided with the set of all Hebrew roots. After a training period, the system is tested for its sensitivity to the location of geminates. To do so, the system must be able to represent the occurrence of bigrams relative to the root boundary. Most importantly, it must distinguish between different types of bigrams: those that contain distinct phonemes vs. geminates. Thus, the system must be able to adequately represent gemination. This requirement appears

simple, but it cannot be met by the pattern associator. Gemination is the copying of a *variable*. Thus, geminates are defined exclusively by their constituency structure. A system cannot ''know'' that the combination *ss* forms a geminate, but *sm* does not, unless it represents their formal structure. But if the representation of constituency structure is rendered obsolete, then how could the system or the child attend to the presence of gemination?

Of course, the answer might be simple: ''They don't! Neither the child nor the system pays any attention to geminates''. This is indeed the only reply available by the pattern associator hypothesis. Geminates must be viewed just like any other bigram, and their acceptability should be predicted by their frequency. In what follows, we examine this proposal by re-analyzing our rating data. We then investigate its adequacy in explaining a sample of evidence drawn from the linguistic literature. The implications of the unique behaviour of geminates are summarized in the final section.

8.2. Are all bigrams created equal? The effect of counter-examples

If geminates are nondistinct from other bigrams, then why is root-initial gemination unacceptable? The answer to this question, according to the pattern association hypothesis, must reside in the distribution of different root tokens in the Hebrew language. We have noted that most Hebrew roots avoid root-initial gemination. Thus, the positional bigram frequency of roots with initial gemination may be lower than that of roots with root-final gemination. The unacceptability of the root [ssm] compared to [mss] would be explained by the rareness of the bigram /ss/ at root initial compared to root final position. Importantly, the structural description of the item should play no part in its rejection. The system is completely blind to the fact that the bigram /ss/ is a geminate. OCP effects are thus viewed as artifacts of statistical structure.

The existence of counter examples to linguistic rules permits dissociating the contribution of their constituency structure from its correlates (Marcus et al., 1995; Pinker, 1991; Prasada and Pinker, 1993). If subjects do not represent gemination, then their ratings for an item should be solely determined by the frequency of its bigrams. The pattern associator hypothesis thus predicts that, if the Hebrew language contained the root [ssm], and this root was highly familiar, then probes manifesting the same initial bigram /ss/ should not be rejected. Conversely, the symbolic hypothesis postulates that the difference between these two root types concerns their structural description. What is 'wrong' with [ssm] is not that /ss/ is a rare bigram but that it is a geminate. The ill-formedness of probes containing a geminate should be maintained regardless of the frequency of their initial bigram.

The structure of Modern Hebrew permits distinguishing the predictions of the two hypotheses. As we have noted, Hebrew manifests several violations of the OCP. Three of these violations are familiar roots in modern Hebrew. Of these roots, two roots, [mmn] and [mmsh] manifest a gemination of the same consonant. These roots are further highly frequent and productive. We are currently pursuing a systematic

investigation of the effect of these counter-examples (Berent et al., 1997). A preliminary evaluation of these issues may be carried out using the present data as well. In our materials, there were two roots formed by initial-gemination of the consonant /m/. Despite being a geminate, the root initial bigram in these items is fairly common. If subjects' ratings are solely determined by the occurrence of specific bigrams then the rejection of items whose initial bigram is common, /mm/, should be blocked. Furthermore, it is expected that these items should be more acceptable compared to the category of the [ssm] type roots as a whole. In contrast, if the rejection of root-initial gemination probes is based on their constituency structure, then these items should nevertheless be unacceptable. They should be rated lower than their controls and similar to other [ssm]-type roots.

To contrast the pattern associator and the symbolic hypotheses, we compared subjects' ratings of the two roots exhibiting root-initial gemination of the consonant /m/ to their root-final gemination and no-gemination controls. We averaged the ratings assigned by each of our subjects across these two items and submitted the means to ANOVA's (3 root type × 3 word class) by subjects. These analyses were conducted separately in each of our studies: Two analyses were conducted on the data from Experiments 1–2, and the third analyses was performed over the replication of Experiment 1, described in the discussion of that study. Mean ratings as a function of root type and word class are presented in Table 4. The main effect of root type was highly significant in each of the these experiments (In Experiment 1: $F(2, 34) = 18.33$, $p = 0.0001$; In Experiment 2: $F(2, 28) = 12.61$, $p = 0.0001$; In the replication: $F(2, 34) = 33.92$, $p = 0.0001$). This main effect was investigated by planned comparisons. In each of the three studies, subjects manifested a significant rejection of root-initial gemination compared to either no-gemination (In Experiment 1: $\Delta = 0.833$, $t(34) = 5.99$, $p = 0.0001$; In Experiment 2, $\Delta = 0.755$, $t(28) = 4.05$, $p = 0.0004$; In the replication: $\Delta = 0.833$, $t(34) = 8.23$, $p = 0.0001$) or root-final gemination (In Experiment 1: $\Delta = 0.5278$, $t(34) = 3.79$, $p = 0.0001$; In Experiment 2, $\Delta = 0.8557$, $t(28) = 4.59$, $p = 0.0001$; In the replication: $\Delta = 0.3797$, $t(34) = 3.753$, $p = 0.0007$). Furthermore, the mean ratings of these items was comparable to that of the [ssm] type root as a whole (see Figs. 3–5).

These findings are incompatible with the predictions of the pattern associator hypothesis. If geminates are indistinguishable from nongeminates, then there is no reason to expect that items whose initial bigram is common should be rejected. The rejection of these items is only explained by their constituency structure, i.e., by the

Table 4
Mean ratings of items generated from roots whose initial bigram is [mm] in Experiment 1, its partial replication, and Experiment 2 as a function of root type and experiment

Root type	Experiment 1	Replication	Experiment 2
[ssm]	1.546	1.593	2.544
[mss]	2.074	1.972	3.399
[psm]	2.380	2.426	3.300

Note: In Experiment 1 and replication: 1 = sounds worst; 3 = sounds best. In Experiment 2: 1 = sounds impossible; 5 = sounds excellent.

fact that their initial bigram, albeit frequent, happens to form a geminate. A failure to distinguish between the representation of geminates and nongeminates cannot account for our subjects' ratings.

8.3. The uniqueness of geminates: Some linguistic evidence

Our preliminary findings regarding the effect of counter-examples suggest that geminates and nongeminate bigrams may not be treated equally. The linguistic literature contains numerous examples supporting the same conclusion (for reviews, see Goldsmith, 1990; Kenstowicz, 1994). In what follows, we briefly describe several illustrations. We show that this evidence could only be accounted for by encoding a structural distinction between geminates and nongeminates.

Consider, for instance, the formation of Hebrew biconsonantal nouns and adjectives from triconsonantal verbs (Aronoff (1994, p. 145); McCarthy (1986, p. 237)). Verbs with triconsonantal roots may form biconsonantal nouns and adjectives from their initial two consonants and the vowel *a*. For example, the verb *DiMeM*, (bled), forms the noun blood, *DaM*. Similarly, the verbs *XiMeM* (heated), *DiLeL* (diluted), and *MeRR* (embittered) form the adjectives *XaM* (hot), *DaL* (diluted, impoverished) and *MaR* (bitter). Interestingly, this generalization applies exclusively to geminates. Verbs whose triconsonantal root does not include a geminate do not form nouns or adjectives in this fashion. Thus, one cannot form adjectives from the verb *BiSheL*, cooked, or *DiBer*, spoke, by means of a biconsonantal root *BaSh or *DaB.

A second illustration of the special behaviour of geminates may be found in antigemination effects, extensively discussed in McCarthy (1986). Phonological rules often avoid gemination. This preference results in blocking phonological processes whose output would yield surface geminates. For instance, syncope rules, (i.e., rules that delete an unstressed vowel) discriminate geminates from nongeminates as a function of their morphemic structure. Hebrew applies syncope rules between the root's nongeminate consonants. For instance, the plural inflection of the singular masculine past tense *KaTaB* is *KaTBu* (wrote). In this form, the root-final consonants are adjacent due to the deletion of the vowel. Similarly, the plurals of *LaMaD*, (learned), and *PaTaR*, (solved) are *LaMDu* and *PaTRu*, each manifesting vowel deletion. However, vowel deletion is blocked if its application yields tautomorphemic geminates. Thus, the plural form for *KaLaL*, (included), is *KaLeLu*, not *KaLLu. Similarly, the plurals of *GaZaZ* (trimmed) and *NaDaD* (wandered) are *GaZeZu* and *NaDeDu*, respectively, not *GaZZu or *NaDDu.

Finally, an essential feature that contrasts geminates with nongeminates is their unalterability. Geminates often block rules whose application would result in the alteration of only one of its members. Kenstowicz (1994) illustrates this property using the following Persian example. Persian realizes the root phoneme [v] as [w] at a syllable's coda. This process is followed by the shifting of [æ] to [o] before [w]. Thus, the root [*nov*] is realized as *nov-i:n* (new) when [v] is in an onset position, but as *now-ru:z* (New Year) when it is at the coda position. However, forms in which the phoneme [v] is part of a geminate block the application of the rule. The integrity of the geminates results in the forms *ævvæl* (first); *morovvæt* (generosity) and *golovv*

(exaggeration), but not *owvæl*, *morowvæt*, or *golowv*. Please note that a geminate [vv] contains the input necessary to trigger the rule, i.e., the phoneme [v]. The blocking of the rule can only be explained by specifying the fact that the bigram [vv] is a geminate. A distinction between identical and nonidentical elements is not limited to the gemination of segments. The sensitivity of linguistic rules to gemination is widely documented and is manifested also with regards to features and tones (Goldsmith, 1990; Kenstowicz, 1994). The ability to distinguish between geminate and nongeminate elements is indispensable for the adequacy of a theory of language.

One may attempt to fix this shortcoming by assuming that bigrams might perhaps be distinguished from nongeminates, but that the distinction need not be structural. After all, bigrams could be viewed as 'very similar' phonemes. So perhaps geminates are just like any other bigram of 'very similar' phonemes: A system that captures the phonetic similarity of forms could handle the behavior of geminates. This solution is not likely to help. In fact, geminates behave just the opposite of what is expected from 'similar' phonemes. McCarthy (1994) documents a statistically significant trend of avoiding adjacent consonants produced by the same articulator in Arabic roots. For instance, Arabic avoids roots like [rfb], which has adjacent labials. Obviously, geminates are produced by the same articulator. If geminates were simply 'similar phonemes', then geminates should have been banned from Arabic roots. This, however, is not the case. Gemination is extremely common, provided it is root-final. Gemination thus cannot be explained by feature similarity. In fact, 'gemination' is not only different than 'phonetic similarity' but appears to exist independently of any specific segmental embodiment. Caramazza and Miceli (1990) report a writing disorder that shifts the location of gemination between distinct phonemes (e.g., *sorella* → *sorrela*). This finding suggests that the notion of gemination is abstract, and independent of the specific segmental contents. It strongly supports the view of geminates as the copying of a variable, rather than the association of tokens.

8.4. Are mental constituents necessary?

The evidence described in the previous sections suggests that subjects discriminate geminates from nongeminate bigrams regardless of their positional frequency. A similar distinction emerges also in numerous other linguistic phenomena in a variety of languages. This distinction, however, may only be expressed by means of constituency structure. The pattern associator hypothesis thus lacks the means to account for gemination and the constraint on its location.

Why is the case of geminates interesting for the study of human cognition, in general, and linguistic knowledge, in particular? After all, a pattern associator mechanism may be fixed to distinguish between geminates and nongeminates. Certainly, an implementation of this distinction in connectionist networks appears simple. For instance, a system may monitor for the presence of geminates (e.g., by enumerating all geminates possible in the language) and devise distinct routines for the processing of geminates and nongeminates. Such implementation may be suffi-

cient to account for our data, although it may be challenged by cases of dissociation between the gemination feature and specific segmental contents (Caramazza and Miceli, 1990).

It is precisely the simplicity of the gemination problem, however, we find insightful. To appreciate its implications it is necessary to distinguish between a mechanism that associates patterns and the pattern associator hypothesis. Our results do not challenge the pattern associator as a computational framework nor do they question the principled adequacy of connectionist models. In fact, some aspects of the OCP phenomenon (e.g., its view as a soft, violable constraint) may be best accommodated in this fashion (see also Goldsmith, 1993; Mohanan, 1993; Prince and Smolensky, 1993). Instead, our argument concerns the pattern associator as a cognitive hypothesis. According to this hypothesis, cognitive processes are blind to formal constituency structure. We showed, however, that a distinction between geminates and nongeminates is necessary to account for linguistic knowledge, and that such a distinction must appeal to constituency structure. A pattern associator mechanism may well be fixed to distinguish between geminates and nongeminates. In doing so, however, the mechanism would violate the pattern association hypothesis. Any solution that differentiates between geminates and nongeminates implements a symbol. Processes whose operation is determined, implicitly or explicitly, by a symbol cannot be said to be insensitive to constituency structure. In fact, such implementations constitute a falsification of the pattern associator hypothesis.

We believe that the essential question for a cognitive theory is not whether constituency structure can be implemented by a pattern-associator. Fodor and Pylyshyn (1988) put that quite clearly: 'of course there are 'sub-symbolic' interactions that implement both rule like and rule violating behavior; for example, quantum mechanical processes do. That's not what classical theorists deny: Indeed, it's not denied by anybody who is even vaguely a materialist.' The ability of connectionist models to implement constituency structure is not questioned. Instead, the debate is whether its implementation is necessary for the adequacy of such systems. The conclusion emerging from the case of geminates suggest that, regardless of their architectural choices, models of linguistic behavior must distinguish between the representations of bigrams by virtue of their internal structure. The uniqueness of geminates, in particular, and the OCP phenomenon, in general, suggests that an adequate account of linguistic generalizations must appeal to the constituency structure of mental representations.

Acknowledgements

References

Aronoff, M. (1994). *Morphology by itself*. MIT Press, Cambridge, MA.

Balota, D., & Chumbley, J. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357.

Bentin, S., Frost, R. (1992). Morphological factors in visual word recognition in Hebrew. In: Feldman, L.B. (Ed.), *Morphological aspects of language processing* (pp. 271–292). Erlbaum, Hillsdale, NJ.

Bentin, S., & Feldman, L.B. (1990). The contribution of morphological and semantic relatedness in visual word recognition: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology*, *42A*, 693–711.

Berent, I., Everett, D. and Shimron, J. (1997). Do phonological representations specify formal variables? Evidence from the Obligatory Contour Principle. Manuscript submitted for publication.

Berent, I., & Perfetti, C.A. (1995). A rose is a REEZ: The two cycles model of phonology assembly in reading English. *Psychological Review*, *102*, 146–184.

Berent, I., Van Orden, G. (1996). Can we unmask the phonemic masking effect? The problem of methodological divergence (pp. 495–500). In: Cottrell, G. (Ed.), *Proceedings of the 18th annual conference of the Cognitive Science Society, Mahwah, NJ.* Erlbaum, Hillsdale, NJ.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*, 297–332.

Caramazza, A., & Miceli, G. (1990). The structure of graphemic representations. *Cognition*, *37*, 243–297.

Emmorey, K. (1995). Processing the dynamic visual-spatial morphology of signed languages. In: Feldman, L.B. (Ed.), *Morphological Aspects of Language Processing* (pp. 29–53). Erlbaum, Hillsdale, NJ.

Even-Shoshan, A. (1993). Ha'milon ha'xadash [The new dictionary]. Kiryat-Sefer, Jerusalem.

Feldman, L.B. (1994). Beyond orthography and phonology: Differences between inflection and derivations. *Journal of Memory and Language*, *33*, 442–470.

Feldman, L.B., & Bentin, S. (1994). Morphological analysis of disrupted morphemes: Evidence from Hebrew. *Quarterly Journal of Experimental Psychology*, *47A*, 407–435.

Feldman, L.B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. Journal of Experimental Psychology: Learning. *Memory and Cognition*, *21*, 947–960.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Fowler, C., Napps, S., & Feldman, L. (1985). Relations among regular and irregular morphologically related words in the lexicon as revealed by repetition priming. *Memory and Cognition*, *13*, 241–255.

Frost, R., Forster, K., Deutsch, A. (1994). Morphological and orthographic analysis in word recognition contrasted. Paper presented at the 35th meeting of the Psychonomic Society, Saint Louis, MO.

Glinert, L. (1989). *The grammar of modern Hebrew*. Cambridge University Press, Cambridge.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Basil Blackwell, Cambridge.

Goldsmith, J. (1993). Harmony phonology. In: Goldsmith J. (Ed.), *The Last Phonological Rule* (pp. 21–60). The University of Chicago Press, Chicago, IL.

Grainger, J., Colé, P., & Segui, J. (1991). Masked morphological priming in visual word recognition. *Journal of Memory and Language*, *30*, 370–384.

Gordon, P. (1989). Levels of affixation in the acquisition of English morphology. *Journal of Memory and Language*, *28*, 519–530.

Greenberg, J. (1950). The patterning of morphemes in Semitic. *Word*, *4*, 196–208.

Kenstowicz, M. (1994). *Phonology in generative grammar*. Blackwell, Cambridge, MA.

Kim, J., Pinker, S., Prince, A., & Prasada, S. (1991). Why no mere mortal has ever flown out to center field. *Cognitive Science*, *15*, 173–218.

Leben, L. (1973). *Suprasegmental phonology*. MIT Press dissertation, Cambridge, MA.

Marcus, G., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, *29*, 189–256.

Mohanan, K. (1993). Fields of attraction in phonology. In: Goldsmith J. (Ed.), *The Last Phonological Rule* (pp. 61–116). University of Chicago Press, Chicago, IL.

Marslen-Wilson, W, Komisarjevsky Tyler, L., & Waksler, R. (1994). Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, 3–33.

McCarthy, J. (1979). *Formal problems in semitic phonology and morphology*. Cambridge, MA: MIT Ph.D. dissertation. Distributed by Indiana University Linguistics Club. Published by Garland Press, New York, 1985.

McCarthy, J. (1986). OCP effects: Gemination and antigemination. *Linguistic Inquiry*, *17*, 207–263.

McCarthy, J.J. (1989). Linear order in phonological representation. *Linguistic Inquiry*, *20*, 71–99.

McCarthy, J. (1994). The phonetics and phonology of Semitic pharyngeals. In: Keating, P. (Ed.), *Phonological structure and phonetic form* (pp. 191–233). Cambridge University Press, Cambridge, MA.

Pinker, S. (1991). Rules of language.. *Science*, *253 (2)*, 530–535.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of parallel distributed processing model of language acquisition. *Cognition*, *28*, 73–193.

Plunkett, K., & Marchman, R. (1993). From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, *58*, 21–69.

Plunkett, K., Nakisa, R.C. (1996). A Connectionist Model of the Arabic Plural System. Manuscript submitted for publication.

Prasada, S., & Pinker, S. (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*, 1–56.

Prince, A., Smolensky, P. (1993). *Optimality Theory*. Rutgers University manuscript.

Rumelhart, D.E., McClelland, J.L. (1986). On learning the past tense of English verbs: Implicit rules or parallel distributed processing? In: McClelland, J.L., Rumelhart, D.E., The PDP Research Group (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 2: Psychological and Biological Models* (pp. 216–271). Bransford Books/MIT Press. Cambridge, MA.

Seidenberg, M. (1987). Sublexical structures in visual word recognition: Access units of orthographic redundancy? In: Coltheart, M. (Ed.), *Attention and Performance XII: Reading* (pp. 245–263). Erlbaum, Hillsdale, NJ.

Seidenberg, M., & McClelland, J. (1989). A distributed developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Stemberger, J.P. (1995). Phonological and lexical constraints on morphological processing. In: Feldman L.B. (Ed.), *Morphological aspects of language processing* (pp. 247–267). Erlbaum, Hillsdale, NJ.

Stolz, J. and Feldman, L.B. (1995). The role of orthographic and semantic transparency of the base morpheme in morphological processing. In: L.B. Feldman (Ed.) *Morphological aspects of language processing* (pp. 109–129). Erlbaum, Hillsdale, NJ.

Stone, G.O., & Van Orden, G.C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception Performance*, *19*, 744–774.

Van Orden, G., Holden, J.C., Podgornik, N.M., and Aitchison, C.S. (1996). When a ROWS is not a ROSE: Null effects and absence of cognitive structures. Manuscript submitted for publication.

Yip, M. (1988). The Obligatory Contour Principle and phonological rules: A loss of identity. *Linguistic Inquiry*, *19*, 65–100.

Yip, M. (1989). Feature geometry and cooccurrence restrictions. *Phonology*, *6*, 349–374.