J. Linguistics **39** (2003), 31–55.
DOI: 10.1017/S0022226702001949 Printed in the United Kingdom

Co-occurrence restrictions on identical consonants in the Hebrew lexicon: are they due to similarity?[1]

IRIS BERENT

Department of Psychology, Florida Atlantic University

JOSEPH SHIMRON

School of Education, University of Haifa

(Received 2 October 2001; revised 7 June 2002)

It is well known that Semitic languages restrict the co-occurrence of identical and homorganic consonants in the root. The IDENTITY HYPOTHESIS attributes this pattern to distinct constraints on identical and nonidentical homorganic consonants (e.g. McCarthy 1986, 1994). Conversely, the SIMILARITY HYPOTHESIS captures these restrictions in terms of a single monotonic ban on perceived similarity (Pierrehumbert 1993; Frisch, Broe & Pierrehumbert 1997). We compare these accounts by examining the acceptability of roots with identical and homorganic consonants at their end. If well-formedness is an inverse, monotonic function of similarity, then roots with identical (fully similar) consonants should be less acceptable than roots with homorganic (partially similar) consonants. Contrary to this prediction, Hebrew speakers prefer root final identity to homorganicity. Our results suggest that speakers encode long-distance identity among root radicals in a manner that is distinct from feature similarity.

It is well known that Semitic grammars constrain the structure of lexically stored forms. To reveal these grammatical constraints, we examine here the predictable regularities in a word's consonantal melody – the sequence of consonants obtained after removing nonreduplicative inflectional affixes and vowels from the word. Because this consonantal melody coincides with the unit listed as 'root' in Semitic dictionaries, this terminology is often maintained even for melodies that exhibit predictable regularities. For instance, the consonantal melody *smm* is dubbed 'root' by sources that clearly argue against its lexical storage in this form (e.g. McCarthy 1981). For the sake of simplicity, we follow here the same tradition. We wish to emphasize, however, that the consonantal melodies we examine are strictly SURFACE forms. We make no claims as to whether these strings are stored as such in the lexicon nor do we argue that they correspond to a phonological or morphological constituent.

[1] We wish to thank Diamandis Gafos, Dan Everett and Janet Pierrehumbert, and two anonymous *JL* referees for helpful discussions of this research.

An inspection of triliteral Semitic roots reveals two patterns of constraints. One concerns the co-occurrence of identical consonants: identical consonants are frequent at the root's end (e.g. *skk*), but rare at its beginning (e.g. *ssk*; Greenberg 1950; Bender & Fulass 1978; McCarthy 1986; Buckley 1997). A second co-occurrence restriction concerns nonidentical homorganic consonants. Nonidentical homorganic root consonants are underrepresented (Greenberg 1950; Bender & Fulass 1978; Pierrehumbert 1993; Buckley 1997; Frisch, Broe & Pierrehumbert 1997). Like identical consonants, homorganic nonidentical consonants are rare root initially (e.g. *kgs*). Unlike identical consonants, however, homorganic nonidentical consonants are infrequent both root initially and root finally (Greenberg 1950).

These distributional patterns have received conflicting explanations in the linguistic literature. According to the IDENTITY HYPOTHESIS, the restrictions on identical consonants are distinct from those concerning nonidentical homorganic consonants (McCarthy 1986, 1994). Conversely, the SIMILARITY HYPOTHESIS offers a single explanation for the restrictions on identical and nonidentical homorganic consonants (Pierrehumbert 1993; Frisch et al. 1997). This view considers identity as a special case of similarity. The restrictions on homorganic (partially similar) and identical (fully similar) segments are therefore captured by a single principle – a statistical monotonic ban on perceived similarity. Although there are various pieces of evidence supporting the similarity hypothesis as an account for the distribution and acceptability of nonidentical homorganic consonants, it is unclear whether this view can specifically handle the restrictions on consonant-identity. The following investigation examines this question. Two experiments compare the acceptability of novel Hebrew roots with identical vs. nonidentical homorganic consonants root finally. If acceptability is an inverse, monotonic function of similarity, then identical (fully similar) segments should be less acceptable than homorganic (partially similar) segments. We show that, despite their greater similarity, identical consonants are in fact MORE acceptable than homorganic nonidentical consonants, a result that is directly opposite to the prediction of the monotonic similarity hypothesis. As we point out, our results are moot with respect to the adequacy of the similarity hypothesis as an account for the co-occurrence of nonidentical homorganic consonants. These results nevertheless favor separate accounts for the co-occurrence of identical and nonidentical homorganic consonants.

1. THE IDENTITY HYPOTHESIS

According to the identity hypothesis, the restrictions on identical consonants are distinct from those affecting nonidentical homorganic consonants. A widely influential version of the identity hypothesis is offered by McCarthy (1986, 1994). McCarthy attributes the restrictions on homorganic and identical consonants to the Obligatory Contour Principle (OCP), a ban on adjacent

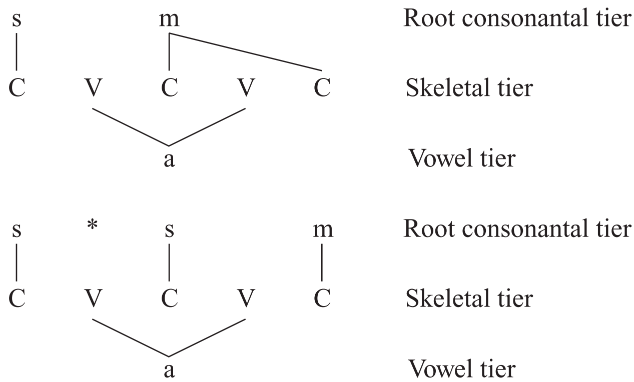

Figure 1
Violations of the OCP-total by roots in the verbs *samam* and *sasam*. Because the root *smm* is formed by spreading, it does not violate the OCP. In contrast, the root *ssm* includes adjacent identical consonants, which violates the OCP-total (OCP violations are indicated by asterisks)

identical elements. Although identical and homorganic consonants are both constrained by the OCP, the target of the OCP in the two cases is different.

The restrictions on identical root consonants reflect the application of the OCP to the entire segment (the root node, hereafter, the OCP-total). The OCP-total thus bans adjacent identical segments. Adjacency, however, is determined by the placement of segments in autosegmental tiers, not by their surface proximity. In particular, because the consonantal root morpheme is represented on a single tier, segregated from vowels and affixes, root consonants are psychologically adjacent, hence, subject to the OCP-total. Accordingly, the OCP-total bans the storage of *ssm*- or *smm*-type forms in the lexicon. Although root geminates may not be stored, they may be formed productively by spreading an underlying biconsonantal representation (e.g. *sm*), as shown in figure 1. The additional assumption that spreading proceeds rightwards correctly predicts the frequency of identical consonants at the end of the root (e.g. *smm*) but not at its beginning (e.g. *ssm*). In contrast, nonidentical homorganic consonants are rare both root initially and root finally (Greenberg 1950). McCarthy (1994) attributes the restrictions on nonidentical homorganic consonants to the application of the OCP at the feature level, specifically, the place node (hereafter, OCP-place). Because homorganic consonants are adjacent on the place of articulation tier, they violate the OCP-place. Furthermore, because, unlike the root node, the place node may not branch to adjacent skeletal positions (McCarthy 1994), adjacent homorganic consonants violate the OCP-place at either initial C1C2 or final C2C3 root positions, as shown in figure 2. The identity hypothesis thus captures the distinct distributional patterns of identical and nonidentical homorganic root consonants by means of separate (albeit related) constraints.

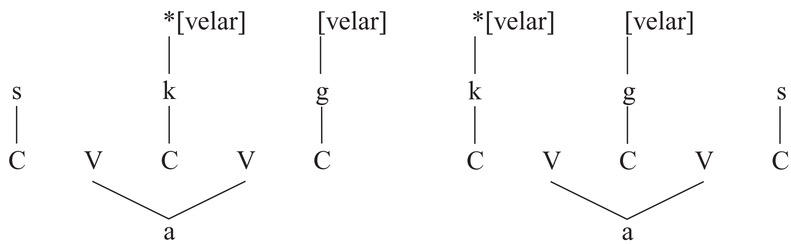

Figure 2
OCP-place violations by the novel verbs *sakag* and *kagas*. Because the velar consonants are adjacent on the place of articulation tier, these verbs each incur violations of OCP-place (OCP-place violations are marked by asterisks)

2. The similarity hypothesis

The similarity hypothesis attributes the restrictions on identical and nonidentical consonants to a single stochastic constraint. On this view, the co-occurrence of consonants depends on their perceived similarity (Pierrehumbert 1993; Frisch et al. 1997). Perceptual similarity, in turn, is defined by two factors: feature overlap and distance. Feature overlap is a weighted function of the number of shared vs. different features between two segments. The distance between segments is not restricted to an underlying representation – Pierrehumbert (1993: 379) notes 'there is no absolute impediment to evaluating the homorganicity of the consonants after vowels are inserted'. Thus, the similarity proposal encompasses two logically distinct claims. One claim is that a stochastic similarity constraint can better capture the co-occurrence restrictions on nonidentical homorganic consonants compared to the OCP-place constraint. A second claim is that the co-occurrence restrictions on identical segments are indistinguishable from those applying to (nonidentical) homorganic consonants.

Consider first the co-occurrence of nonidentical homorganic consonants. Pierrehumbert (1993) argued that the co-occurrence of homorganic root consonants is inexplicable by a categorical constraint on place of articulation, as their distribution is further sensitive to secondary place of articulation and noncontrastive voicing (see also Frisch et al. 1997). Additional criticisms of the OCP-place concerns locality. The OCP bans adjacent homorganic segments. Because adjacency is defined relative to autosegmental representations that encode homorganic consonants on a single tier, homorganicity effects should be blind to intervening elements that are not specified for that place of articulation. For instance, the co-occurrence of the coronal obstruents *td* should be unaffected by the intervening labial *b*. The OCP-place thus predicts similar-size dissimilation effects for adjacent (e.g. ***tdb***) and nonadjacent (e.g. ***tbd***) homorganic consonants. Contrary to this prediction, however, the observed/expected ratio for homorganic consonants is lower for adjacent consonants relative to nonadjacent consonants (Pierrehumbert 1993; Buckley 1997;

Frisch et al. 1997). These observations suggest that the constraint on nonidentical root consonants is not limited to adjacent root radicals, nor is it specific to homorganicity. Pierrehumbert (1993) proposed to capture the co-occurrence of nonidentical root radicals by a stochastic constraint on perceived similarity. Frisch et al. (1997) extended this proposal by providing a detailed formal analysis of perceived similarity. They further demonstrated, by means of a regression analysis, that a structured specification account of similarity provides a superior fit for the distributional data compared to other proposals, including McCarthy's categorical account, a frequency account, and Pierrehumbert's (1993) proposal, which captures similarity in terms of contrastive underspecification.

The stochastic account, however, not only offers an alternative for the OCP-place, but further aspires to subsume the OCP-total. Similarity is presented as a single unified explanation for the co-occurrence of homorganic (nonidentical) and identical consonants. Identical consonants, on this view, are consonants that are maximally similar. The co-occurrence of identical consonants thus presents the limiting case of a more general stochastic constraint on perceived similarity. In support of this view, Pierrehumbert (1993), Frisch et al. (1997) and Buckley (1997) demonstrate that the distribution of identical consonants is constrained by perceived similarity and distance, the same principles used to capture the distribution of nonidentical homorganic consonants. For instance, in accord with the view of identical elements as maximally similar, there is a larger gap in the co-occurrence of nonadjacent identical consonants (e.g. *tbt*) compared to homorganic nonidentical consonants (e.g. *tbd*). This finding is unexpected on McCarthy's (1994) account: because the OCP-total does not operate across an intervening C2 segment, *tbt* and *tbd* should have incurred a comparable violation of the OCP-place (Pierrehumbert 1993). Identical and homorganic nonidentical consonants alike are further affected by distance: they are less likely to occur in adjacent (C1C2 and C2C3) compared to nonadjacent (C1C3) positions; and the magnitude of the distance effect depends on similarity: nonadjacent homorganic consonants are more frequent than nonadjacent identical consonants. None of these findings is expected on McCarthy's account. Pierrehumbert (1993) and Frisch et al. (1997) thus conclude that the distribution of identical and homorganic nonidentical root consonants is governed by a single constraint on perceived similarity.

3. Some challenges to the similarity account

Although there is some compelling evidence favoring the similarity hypothesis as an account for the distribution of homorganic root segments, these observations cannot unequivocally support its potential for capturing the distribution of identical root radicals. The main problem is the failure to fully account for the distribution of identical consonants. According to the similarity hypothesis, the acceptability of root radicals is a monotonic function of their perceived similarity. Because identical consonants are maximally similar, they should be less acceptable than nonidentical consonants. This prediction could be evaluated at either root initial or root final position. The evaluation of this prediction for root initial consonants is not entirely clear: because root initial identical consonants violate the OCP-total, their greater unacceptability may be due either to their greater similarity or to multiple constraint violation (i.e. the violation of OCP-place and OCP-total). A greater unacceptability of identical and homorganic consonants at root initial position is thus consistent with either the similarity or the identity hypotheses. The predictions of these two views diverge, however, for root final consonants. The monotonic similarity account predicts lower acceptability for identical (fully similar) relative to homorganic (partially similar) consonants. Conversely, the identity hypothesis allows for the constraints on identical consonants to differ from those governing similar consonants. In fact, the predictions of McCarthy's (1986) account are clearly contradictory to the similarity hypothesis. Because *smm*-type roots do not violate either the OCP-total or the OCP-place, such roots should be clearly more acceptable than roots with final homorganicity. Root final identity thus allows adjudicating between the similarity and identity hypotheses. Unfortunately, existing tests of the similarity hypothesis do not consider the distribution of *smm*-type roots. Recall that the similarity hypothesis is supported by the observation of gaps between the observed and expected co-occurrence of root radicals. These analyses, however, exclude roots with final identical consonants on the grounds that these are lexically represented in their biconsonantal form, as proposed by McCarthy (1986).

On superficial examination, this decision appears unproblematic: if 'root' is a unit of lexical storage, and if lexical storage is contingent on identity erasure, then *smm*-type forms are not truly roots. Accordingly, the distribution of *smm*-type forms falls beyond the scope of an account of the lexicon. We see several problems with this reasoning, however. To be sure, we do not dispute the claim that verbs such as *simem* are subject to identity erasure, nor do we argue that they are lexically stored as *smm*. As explained above, we use the term 'root' to refer to a surface unit – the string of consonants obtained after removing nonreduplicative inflectional affixes and vowels. This unit may not necessarily correspond to the unit of lexical storage, inferred after applying grammatical constraints to surface representations. Nevertheless, we believe that the exclusion of *smm*-type forms limits the evaluation of the similarity hypothesis. First, the exclusion of *smm*-type roots on the grounds that they are not lexically stored tacitly confines the similarity hypothesis to the lexicon. This assumption does not follow from the similarity hypothesis. In fact, proponents of this proposal specifically argue against such an approach (Pierrehumbert 1993; Frisch et al. 1997). Second, the exclusion of *smm*-type forms remains unjustified even if the lexical domain is explicitly assumed.

A principal motivation for the similarity hypothesis is the postulation of a single explanation for the distribution of identical and homorganic root consonants. Specifically, Pierrehumbert (1993: 373) notes that

> [i]n the standard model, the gradience of the dissimilarity requirements in the Arabic verbal roots is disregarded, and the cooperation of several grammatical mechanisms is invoked to cover those regularities which are described. In the alternative approach to be developed here, the gradience is exploited both to improve the descriptive coverage and to unify the most categorical effect (the total OCP on adjacent consonants) with less categorical effects. … We propose that Arabic has an OCP effect of place alone. It applies only to consonants which are perceived to be similar, and the strength of the effect increases with perceived similarity. Identical consonants are viewed as maximally similar, and so the total OCP arises as the limiting case of maximally strong enforcement of the place OCP.

Disregarding *smm*-type forms is inconsistent with this goal. If the similarity hypothesis maintains that the distribution of lexical forms is fully explicable by their perceived similarity, irrespective of identity, then it cannot simultaneously maintain that identity erasure is operative in the lexicon. It is the theory's task to explain how the language learner comes to acquire such a lexical representation. There are two possible replies to this question. One is that the erasure of the final radical from *smm*-type roots is due to C2C3 identity. This assumption indeed lies at the core of McCarthy's (1986) OCP-total proposal – an account that presupposes an innate grammatical constraint on segment identity. The identity hypothesis, however, is not limited to McCarthy's (1986) OCP-total: it includes any proposal that constrains segment identity, regardless of whether it is violable or inviolable, innate or learnable. The similarity account is clearly free to adopt any of these versions of identity-avoidance, but this move would EMBRACE the identity hypothesis, not subsume it. The other possibility is that the erasure of the final radical from *smm*-type forms is due to C2C3 similarity, not specifically their identity. This is indeed the only logical possibility available if the identity hypothesis is rejected. The challenge to this account is to explain why root radicals are routinely erased when they are fully similar, but never when they are only partly similar. In the absence of an independent explanation for identity erasure, the exclusion of *smm*-type roots from the lexicon remains unmotivated.

The unmotivated exclusion of *smm*-type roots indeed compromises the main empirical support for the similarity hypothesis, namely, its ability to capture the under-representation of identity in the lexicons of Arabic and Tigrinya. Indeed, if one were to apply these calculations to all triliteral roots – not just to hypothesized lexical representations – then the observed frequency of roots with identical elements would be likely to exceed that of roots having homorganic elements at root final position (Greenberg 1950; Buckley 1997).

	C1C2		C2C3	
	Identical	Homorganic	Identical	Homorganic
Coronal sonorants (l, r, n)	0	1	38	2
Labials (b, p, m, v)	2	2	27	0
Velars (k, g)	1	0	17	0
Gutturals (ʔ, ʕ, h, x)	0	8	12	17
Coronal Obstruents				
stops (d, t)	1	0	23	1
fricatives (z, s, ʃ)	0	3	21	0
affricates (c)	0	0	7	0

Table 1

The distribution of identical and homorganic root consonants in Hebrew at initial (C1C2) and final (C2C3) positions

For Hebrew, our own count of the set of 1449 productive roots listed in the Even-Shoshan (1993) dictionary corroborates these conclusions.[2] Specifically, whereas root initially nonidentical homorganic consonants are somewhat more frequent than identical consonants, at root final positions the pattern is reversed:[3] root final identity is generally more frequent than homorganicity

[2] Because of the historical changes in the Hebrew phonemic system, the classification of several phonemes for place of articulation is uncertain. One case in point concerns the class of gutturals. Biblical Hebrew included four gutturals: the pharyngeal fricatives /ħ/ and /ʕ/, the glottal stop /ʔ/, and the glottal fricative /h/. Modern Hebrew has lost the phonemes /ħ/ and /ʕ/. The Biblical /ʕ/ and /ħ/ are realized as /ʔ/ and /x/, respectively, by most contemporary speakers. Although the contemporary [x] is a velar, it nevertheless retains some of the phonemic characteristics of a guttural (Bolozky 1978; Sandler 1994), and is therefore considered as such for the purposes of the present analysis. Another phonemic distinction lost among most speakers of Modern Hebrew is the emphatic consonants /t̥/, /s̥/ and /q/, which are realized as /t/, /c/ and /k/, respectively (Bolozky 1978). Although the distinction between these emphatic consonants and their nonemphatic realizations is maintained in the orthographic transcription given in Even-Shoshan (1993), it does not affect the calculation of consonant co-occurrence, as the emphatic and nonemphatic members never co-occur in adjacent positions (C1C2 or C2C3). Hebrew also has a highly predictable spirantization rule that realizes /b/ and /p/ as [v] and [f], respectively (Bolozky 1978). Our analysis ignores these predictable alternations and considers only the underlying representation, /b/ and /p/. Finally, the homorganicity of a consonant combination with the affricate /c/ is computed under the assumption that its manner is represented as a stop-fricative sequence. Accordingly, affricate-fricative and stop-affricate sequences are considered homorganic. Our analysis does not report the distribution of the phoneme /y/, as it is the only Hebrew glide.

[3] The one exception to the rule is the overrepresentation of homorganic relative to identical gutturals at C2C3 positions. This observation, however, appears to be entirely due to the frequency of the radical *h* at the root's end in the Even-Shoshan (1993) dictionary. Even-Shoshan (1993) uses the radical *h* to capture the root final radicals of weak roots. This transcription, however, reflects an orthographic convention, as the *h* radical is never realized in the stem. To maintain consistency with our source, we coded these roots as listed there. The

(see table 1). This observation is contrary to the predictions of the similarity hypothesis.

Although the discrepancy between the predicted and the observed distribution of root final identity is unexpected under the similarity hypothesis, it does not necessarily falsify this view. The similarity hypothesis captures speakers' knowledge of root phonotactics. This knowledge, however, may not be directly evidenced in the distribution of lexical forms. Distributional facts may be shaped by nonlinguistic or diachronic factors that are not currently active in a language; hence they may not necessarily reflect the linguistic competence of modern speakers (Everett & Berent 1998). Experimental work allows for additional insights into speakers' synchronic knowledge of root structure. We now proceed to discuss the existing results.

4. Behavioral evidence for co-occurrence restrictions on root structure

4.1 *Co-occurrence restrictions on root identity in Hebrew*

Recent experimental evidence demonstrates that speakers of Modern Hebrew are sensitive to the co-occurrence of identical root consonants. For instance, novel Hebrew roots with identical consonants at their beginning (e.g. *ssm*) are rated less acceptable than roots with adjacent identical consonants at their end (e.g. *smm*, Berent & Shimron 1997; Berent, Everett & Shimron 2001(a)). The sensitivity to the location of identical root radicals is also observed online. Consider the lexical decision task, for instance. Participants in this task are asked to quickly indicate whether or not a string of letters corresponds to an existing word by pressing one of two computer keys. The speed and accuracy of classifying a stimulus may be used to gauge its well-formedness: novel words that violate co-occurrence restrictions should be classified as nonwords more easily than novel words that do not violate such restrictions. Berent and colleagues found that lexical decision is specifically sensitive to the location of identical consonants in the root: novel words with *ssm*-type roots were classified as nonwords faster and more accurately than *smm*-type controls, a finding consistent with the view of *ssm*-type roots as ill-formed (Berent, Shimron & Vaknin 2001(b); Berent, Marcus, Shimron & Gafos 2002). The ill-formedness of *ssm*-type roots is evident even when no explicit response to the word is required, as evidenced in the Stroop task (Berent, Bibi & Tzelgov 2000). Here, participants are asked to name the color of letter strings printed in color (e.g. the word *sisem* printed in red) while ignoring the contents of the printed word. The speed of color naming (e.g. saying 'red') is used to gauge the internal ill-formedness of the root: because ill-formed stimuli are easier to

ubiquity of root final *h* with homorganic radicals, however, should not be considered as evidence for weakening of the OCP-place for root final gutturals.

ignore, they are expected to interfere with color naming less than well-formed stimuli. In accord with the view that root initial gemination is ill-formed, Berent et al. (2000) observed faster color naming latencies for words derived from the *ssm*-type compared to *smm*-type controls. The sensitivity to the root location of identical consonants in online tasks, including tasks that do not require attention to root structure, suggests that it forms part of speakers' linguistic competence.

Existing experimental investigations can further rule out several alternative explanations for the co-occurrence restrictions on identical root consonants. One explanation attributes the experimental results to a restriction on word, rather than root structure. For instance, the ill-formedness of *ssm*-type roots could be due to the ill-formedness of C_1VC_1 combinations word initially (Frisch & Zawaydeh 2001). The experiments reported above systematically examined this account by manipulating the location of the root within the word. The findings consistently revealed sensitivity to the location of identical consonants across various word positions and despite intermediate vowels and affixes. These results suggest that the restrictions on the location of identical consonants are defined relative to a morphological constituent (the root or stem), rather than to their word position. A second alternative explanation attributes the restrictions on identical consonants to the distribution of specific root tokens in speakers' lexicons. On this account, the unacceptability of the root *ssm*, for instance, is due to the rarity of its consonant-combinations (especially the initial geminates, e.g. *ss*), rather than to the violation of an abstract constraint on identity, a formal relationship between any consonant pair (e.g. *XX, where X stands for any consonant). Contrary to this view, the unacceptability of *ssm*-type roots was observed even when *ssm*- and *smm*-type roots were equated for familiarity. Specifically, Berent et al. (2002) observed the asymmetry in the location of identical consonants for roots including novel geminate phonemes; hence *ssm* and *smm*-type roots were equally (un)familiar. Likewise, speakers discriminate between novel *smm*-type roots and no-gemination controls where the materials are equated for the co-occurrence of root radicals (Berent et al. 2001(a)). These results demonstrate that speakers are sensitive to root structure, and that they can discriminate between various root types depending on the presence of identical consonants and their location in the root.

4.2 *The source of the restrictions on identical consonants: identity vs. homorganicity*

Although the existing empirical findings clearly demonstrate sensitivity to the location of identical consonants in the root, they cannot unequivocally reveal its source. In particular, the existing results cannot determine whether this co-occurrence restriction specifically concerns identity or homorganicity. Recent experiments by Frisch & Zawaydeh (2001) report that speakers of Jordanian

Arabic consider novel CaCaCa verbs with homorganic consonants as less acceptable than verbs with no homorganic consonants. Although the generality of homorganicity avoidance across morphological patterns remains to be seen, these results suggest that speakers disfavor root homorganicity. If nonidentical, homorganic consonants are actively disfavored, then it is possible that, despite the greater frequency of roots with final identical consonants in the lexicon (relative to homorganic consonants), identity may not be fully desirable. In fact, Hebrew roots with final identity are rated as less acceptable than no-gemination controls. For instance, in Berent et al.'s (2001(a)) experiment 2, roots with final identity (e.g. *smm*) were assigned a mean acceptability rating of 2.02, whereas the mean for roots with no identity was 2.58 (using a 1–3 scale, 1 = sounds best, 3 = sounds worst). Although *smm*-type roots are certainly acceptable, their ratings were consistently lower than roots with no identity (for similar findings, see also Berent & Shimron 1997). The relative unacceptability of *smm*-type roots could be due to their similarity. More generally, the acceptability of Semitic roots may be captured by a single monotonic function of their similarity without assuming separate restrictions on identity. Although Frisch & Zawaydeh (2001) have raised some doubts regarding the adequacy of the similarity hypothesis as an account for the restrictions on root-identity (see also Frisch in press),[4] the similarity account for the acceptability of root identity has not been systematically investigated. In what follows, we examine the monotonic similarity hypothesis experimentally by comparing the acceptability of roots including homorganic and identical consonants. If acceptability is an inverse, monotonic function of similarity, then identical (fully similar) segments should be less acceptable than homorganic (partially similar) segments.

5. Experimental design and predictions

The following experiments compare the acceptability of roots with identical and homorganic, nonidentical consonants. The view of acceptability as a monotonic, inverse function of similarity predicts that identical consonants should be less acceptable than nonidentical homorganic consonants. Although the lesser acceptability of identical compared to nonidentical homorganic consonants is predicted at both root initial and root final

[4] Frisch & Zawaydeh (2001) reported higher ratings for verbs with identical root final consonants compared to verbs with homorganic root consonants. This post-hoc observation, however, was based on only three root tokens and was not evaluated statistically. Furthermore, because these authors did not systematically control for the location of homorganic consonants in the root or the word, the observed trend may be due to either of these factors. For instance, because homorganic consonants may be either root final or root initial, the preference for identical root final consonants over homorganic consonants may be explained by a preference concerning the location of homorganic consonants in the root (a preference for final over initial homorganic consonant), rather than a distinction between identity and homorganicity.

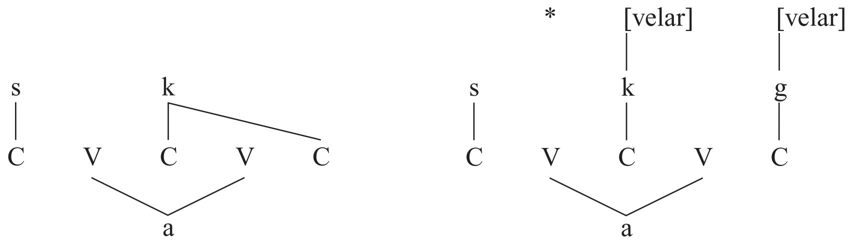

Figure 3
OCP violations by homorganic vs. identical c2c3 consonants for the novel roots *skk* and *skg*. Identical consonants in the root *skk* incur no OCP violations, whereas the homorganic consonants in *sgk* violate the OCP-place

positions, the interpretation of this effect is clearest root finally. As explained above (see section 3), a greater acceptability of *ssm*- relative to *szm*-type roots is explicable by either similarity or multiple constraint violation (OCP-place and OCP-total). A greater unacceptability of identical compared to homorganic consonants at root initial position is thus consistent with both the similarity and the identity hypotheses. The predictions of these two views differ, however, for root final consonants. The monotonic similarity account predicts lower acceptability for identical (fully similar) relative to homorganic (partially similar) consonants. Conversely, the identity hypothesis makes it possible to dissociate the acceptability of identical consonants from their similarity. Specifically, on McCarthy's (1986) version of the identity hypothesis, *smm*-type roots do not violate either the OCP-place or OCP-total. Accordingly, *smm*-type roots should be more acceptable compared to roots with C2C3 homorganic consonants, which violate OCP-place (see figure 3). Roots with final identity thus make it possible to test whether the acceptability of identical consonants is explicable by their similarity. Our experimental investigation compares roots with identical and homorganic (nonidentical) consonants at their final positions.

Our experiments elicit acceptability ratings for root trios comprised of three members. One member had a root with adjacent nonidentical homorganic consonants at the root's end (e.g. *nkg*), a second member had adjacent identical consonants at the root's end (e.g. *skk*), and the third member had no homorganic root radicals (e.g. *nks*; see table 2). To assure that the distinction between root types was not due to familiarity with specific consonant combinations, we evaluated the three root types (homorganic, identical and nonhomorganic controls) for their radical co-occurrence in the Hebrew lexicon. To further secure the attribution of acceptability ratings to restrictions on the structure of roots, rather than words, we systematically varied the position of these roots within the word by conjugating each root in three classes of word patterns. If the restrictions on homorganic and identical consonants truly

Root		Word class FIRST Unaffixed	SECOND Affixed adjacent	THIRD Affixed nonadjacent
Homorganic	nkg	nikeg	mankigot	hitnakagta
Identical	skk	sikek	maskikot	histakakta
Control	nks	nikes	mankisot	hitnakasta

Table 2

The structure of the experimental materials used in experiments 1 and 2

concern their position in the root, then these restrictions should be observed regardless of word position and irrespective of intermediate vowels. The first class corresponded to unaffixed words in the verbal patterns known as BINYANIM PIʔEL and KAL, singular masculine past. The location of the C2C3 radicals in these verbs was invariably word final. In contrast, members of the other two classes were preceded by a prefix and followed by a suffix; hence, the location of the root's C2C3 radicals was word internal. The second word class included the nominal patterns known as MISHKALIM MAFʔIL and MIFʔAL, whereas the third word class included verbs in BINYAN HITPAʔEL past tense. Each of the three root members was conjugated in precisely the same word pattern, and was presented in each of the three classes. Because the resulting words were perfectly matched on all aspects other than root structure, acceptability differences between trio members should uniquely reflect the effect of root structure. We obtained two separate acceptability measures. Experiment 1 compared the acceptability of members of the root trios relative to each other (hereafter: relative rating), whereas experiment 2 examined the acceptability of each member on its own (hereafter: absolute rating).

All accounts predict lower acceptability for roots with homorganic consonants compared to controls. Of primary interest is the comparison of identical and nonidentical homorganic consonants. If the co-occurrence of homorganic and identical consonants is a monotonic function of their perceived similarity, then roots with identical consonants (e.g. *skk*) should be considered less acceptable compared to roots with homorganic, nonidentical consonants (e.g. *nkg*). The predictions under the identity hypothesis are reversed: because roots with final identity do not violate either the OCP-place or OCP-total, they should be more acceptable compared to roots with homorganic, nonidentical consonants. If the restrictions on homorganic and identical consonants concern their root position, then the same outcome should emerge across word classes, regardless of the position of the root in the word.

6. Method

6.1 *Participants*

Participants were University of Haifa students who were native Hebrew speakers. They were tested in two groups (for experiments 1 and 2, respectively) in a class setting and received no compensation for participating in the experiment. Twenty-four participants took part in experiment 1. An additional group of twenty-two students participated in experiment 2.

6.2 *Materials*

The materials consisted of novel words, generated by conjugating 13 novel roots in existing word templates. The roots were arranged in trios. One member of the trio had adjacent homorganic consonants root finally, the second root member had adjacent identical consonants root finally and a third member included no homorganic consonants. The homorganic and identical trio members were matched on the place of articulation of the homorganic consonant pairs. Homorganic consonants were sampled from one of five natural classes: nonsonorant coronals (t, z, ʃ, c; a total of four trios), labials (b, m, p; a total of four trios), velar stops (g, k; a total of two trios), gutturals (x, ʔ, ʕ; a total of two trios) and coronal sonorants (n, l; a total of one trio).[5]

To assure that the distinction between members of the trio was not due to familiarity with specific consonant combinations, we matched the three types of roots for the co-occurrence of their radicals. Our calculation of the co-occurrence of root radicals was based on a database of 1449 productive tri-consonantal roots listed in the Even-Shoshan dictionary (1993). We consider a root productive if it appears in at least one verbal form. The database lists the orthographic transcription of all the triliteral productive roots in Even-Shoshan (1993), including *smm*-type roots. Using this database, we determined the frequency of any two root-radicals (bigrams), including both adjacent (i.e. C1C2, C2C3) and nonadjacent (i.e. C1C3) combinations. For each member of the trio, we calculated the sum positional bigram frequency. For instance, the novel root *gkl* has a summed bigram frequency of 14, because its C2C3 combination occurs in four roots (*skl*, *tkl*, *ʃkl*, *ʕkl*), its C1C3 bigram occurs in ten roots (*grl*, *gll*, *gʔl*, *gʕl*, *gyl*, *gbl*, *gdl*, *gml*, *gzl*, *gxl*) and its C1C2 bigram does not occur in any root. The mean summed positional bigram frequencies were 11.5 (SD = 2.0), 12.5 (SD = 2.5) and 11.8 (SD = 2.3) for

[5] The experiment also included three additional trios with the phoneme /r/. Many speakers of Modern Hebrew encode this consonant as the uvular fricative /ʁ/ (Sandler 1994). In view of the great variability in the pronunciation of this phoneme, its classification for homorganicity is uncertain. Accordingly, we excluded these items from all subsequent analyses.

the homorganic, identical and control roots, respectively. None of these mean-differences approached significance by planned comparisons.

The experimental materials were generated by conjugating each root trio in one of three classes of word patterns, as described above. Note that the conjugation of roots could potentially lead to OCP violations due to identity or homorganicity between the stem and suffixes. This problem is particularly prevalent in the verbal system, which includes the suffixes *ti*, *ta*, *t*, *nu*, *tem*, *ten*. To minimize such effects, we refrained from concatenating suffixes that share the same manner and place of articulation with the root final radical. Accordingly, roots ending with a coronal sonorant consonant were not paired with the suffix -*nu*. Likewise, roots ending with a coronal obstruent stop were never followed by suffixes beginning with a coronal obstruent stop. Furthermore, because roots with identical and homorganic C2C3 were matched for place of articulation and suffix, any residual effects of suffix-stem similarity must apply equally to both pair members. Accordingly, the differences between identical and homorganic roots must reflect their root structure. Each root trio was conjugated in each of these three word classes, resulting in a total of 39 word trios.

6.3 *Procedure*

Participants were presented with a printed list including 117 words, generated by conjugating the 13 root trios described above in three word classes. The words were typed and their vowels were encoded by the standard diacritic marks. Participants were instructed to read each word carefully, while attending to its pronunciation, and rate its acceptability as a possible Hebrew word. The two rating experiments differed, however, in the precise rating judgment that was required.

Participants in the relative rating procedure (experiment 1) were presented with the experimental materials typed in trios. The order of the word trios was random. Likewise, the order of the words within each word trio was randomly determined. They were next instructed to compare the words in the trio and assess the extent to which they sounded like a possible Hebrew word. Participants were asked to assign the rating 1 to the word that sounded the best, 2 to the word that sounded intermediate, and 3 to the word that sounded the worst. To express high acceptability ratings by larger numbers, we inverted the scale by subtracting each score from 4. Thus, in our report, 1 corresponds to the word that sounds worst and 3 indicates the word that sounds best.

In the absolute rating experiment, the same words were presented in a single randomized list, and participants were asked to provide a rating for each word in isolation (rather than compare it to its matched trio members). Responses were provided using a scale of 1–5 (1 = worst, 5 = best).

7. RESULTS

7.1 *Relative rating*

Acceptability ratings were submitted to 2-way ANOVAs by participants and items. All results referred to as significant have a p-value that is lower than 0.05. Both of these analyses yielded a significant effect of root type ($F_1(2, 46) = 54.79$, $MSE = 0.116$; $F_2(2, 24) = 7.03$, $MSE = 0.497$). The main effect of word class ($F_1(2, 46) = 4.60$, $MSE = 0.001$; $F_2(2, 24) < 1$, $MSE = 0.000$, n.s.) and its interaction with root type ($F_1(4, 92) = 4.00$, $MSE = 0.057$; $F_2(4, 48) = 1.32$, $MSE = 0.098$) were significant in the analysis by participants but not by items ('n.s.' is used throughout to indicate nonsignificant effects). Figure 4 plots mean acceptability ratings as a function of root type and word class.

The predictions of the similarity and identity hypotheses were further tested using planned comparisons. Both accounts predict that homorganic roots should be less acceptable than nonhomorganic controls. This prediction was supported across word classes ($F_1(1, 46) = 109.29$; $F_2(1, 24) = 14.05$), as well as within each of the three word classes separately (in the first word class: ($F_1(1, 92) = 49.54$; $F_2(1, 48) = 19.73$); in the second word class: ($F_1(1, 92) = 100.45$; $F_2(1, 48) = 33.33$); in the third word class: ($F_1(1, 92) = 62.56$; $F_2(1, 48) = 19.51$)). The critical prediction that contrasts the two accounts, however, concerns roots with homorganic vs. identical consonants. Contrary to the prediction of the monotonic similarity hypothesis, roots with identical consonants were more acceptable than the less similar, homorganic roots. This trend reached significance in the analysis by participants across word classes ($F_1(1, 92) = 22.69$; $F_2(1, 48) = 3.12$, $p = 0.09$). Likewise, identical consonants were significantly more acceptable compared to homorganic consonants in the first ($F_1(1, 92) = 26.96$; $F_2(1, 48) = 9.35$) and third word classes ($F_1(1, 92) = 19.96$; $F_2(1, 48) = 6.16$), as well as in the analysis by participants (but not by items) in the second word class ($F_1(1, 92) = 4.46$; $F_2(1, 48) = 1.84$, $p = 0.18$, n.s.).

The greater acceptability of fully similar, identical root radicals compared to partially similar, homorganic radicals is incompatible with the view that root structure is governed by a single monotonic constraint on perceived similarity. Instead, this finding suggests that the constraint on total identity is distinct from the restriction on place of articulation. Our findings, however, suggest that root final identity is not entirely acceptable. Roots with identical consonants were rated as less acceptable than controls across word classes ($F_1(1, 46) = 32.39$; $F_2(1, 24) = 3.93$, $p = 0.06$). This trend was significant in the second ($F_1(1, 92) = 62.56$; $F_2(1, 48) = 19.50$) and it was significant by participants in the third ($F_1(1, 92) = 11.84$; $F_2(1, 48) = 3.74$, $p = 0.06$) and first word classes ($F_1(1, 92) = 7.38$; $F_2(1, 48) = 1.91$, $p = 0.18$, n.s.). Note that the unacceptability of roots with identical consonants cannot be explained simply by their homorganicity, as these roots were rated significantly higher

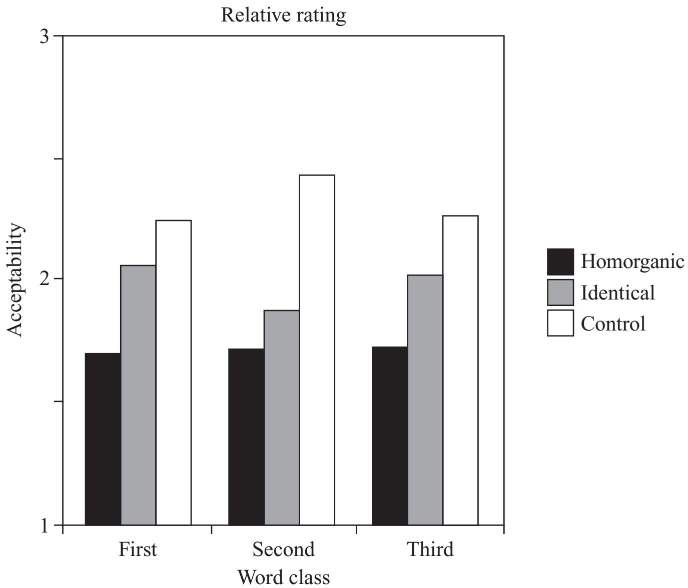

Figure 4
Mean acceptability ratings for roots with final homorganic consonants, final identical consonants and controls as a function of root type and word class (experiment 1)

than the homorganic roots. We return to discuss the reasons for the partial unacceptability of *smm*-type roots in section 8.

7.2 *Results – absolute rating*

Experiment 2 presents a more stringent test for the identity hypothesis by introducing a simple modification to the rating procedure. Participants in experiment 2 were presented with the same words as were used in experiment 1. These words, however, were now arranged in a single randomized list, rather than in matched trios. Participants were simply asked to determine the acceptability of each word on its own (instead of comparing it to its matched trio members). The change in the rating procedure makes it possible to assess the generality of the restrictions on identical consonants. Indeed, one may be concerned that the explicit demand to compare words that differ solely in their root structure (in experiment 1) could give rise to deliberate, nonlinguistic discrimination strategies. In contrast, the absolute rating task does not call attention to root structure, as words may be distinguished by a variety of other factors, such as their word class, the place of articulation of root consonants, etc. The absolute rating task thus minimizes the role of deliberate, nonlinguistic discrimination strategies relative to the comparative rating task used

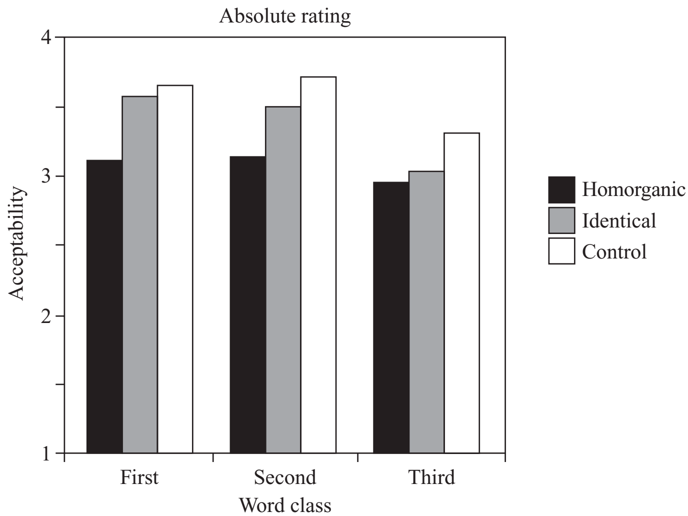

Figure 5
Mean acceptability ratings for roots with final homorganic consonants, final identical consonants and controls as a function of root type and word class (experiment 2)

in experiment 1. If the restrictions on identical consonants form part of speakers' linguistic competence, and are routinely applied in the representation of novel words, then we expect to observe a distinction between identical and homorganic consonants even when attention to root structure is not highlighted.

The results of this experiment offer support for the identity hypothesis. Participants' ratings were analyzed by ANOVAs (3 root types × 3 word classes) using both participants and items as random variables. The ANOVAs yielded a significant main effect of root type ($F_1(2, 42) = 33.14$, $MSE = 0.121$; $F_2(2, 24) = 7.65$, $MSE = 0.319$) and word class ($F_1(2, 42) = 4.77$, $MSE = 0.55$; $F_2(2, 24) = 11.61$, $MSE = 0.132$), and a marginally significant interaction of root type × word class ($F_1(4, 84) = 2.42$, $MSE = 0.087$; $F_2(4, 48) = 1.86$, $MSE = 0.068$, $p = 0.13$, n.s.). Figure 5 plots mean acceptability ratings as a function of root type and word class.

Planned comparisons demonstrated that, across word classes, roots with homorganic consonants were rated as less acceptable than controls, a finding that emerged across word classes ($F_1(1, 42) = 65.18$; $F_2(1, 24) = 15.06$) as well as within each of the word classes separately (in the first word class: ($F_1(1, 84) = 36.53$; $F_2(1, 48) = 27.90$); in the second word class: ($F_1(1, 84) = 42.99$; $F_2(1, 48) = 33.18$); in the third word class: ($F_1(1, 84) = 15.57$, $F_2(1, 48) = 12.09$)). Replicating the findings of experiment 1, however, roots with identical consonants root finally were significantly more acceptable than roots with homorganic consonants ($F_1(1, 42) = 24.49$; $F_2(1, 24) = 5.57$). This finding was significant for the first ($F_1(1, 84) = 25.20$; $F_2(1, 48) = 19.63$) and second

($F_1(1, 84) = 16.79$; $F_2(1, 48) = 12.09$) word classes, but not for the third word class ($F_1(1, 84) = 1.05$, $p = 0.31$, n.s.; $F_2(1, 48) < 1$, $p < 0.3$, n.s.). Although roots with identical consonants were more acceptable than homorganic roots, they were not considered perfectly acceptable. Across word classes, roots with final identical consonants were rated less acceptable than controls, a finding that was significant by participants but not by items ($F_1(1, 42) = 9.76$; $F_2(1, 24) = 2.31$, $p = 0.14$, n.s.). The relative unacceptability of *smm*-type roots was significant in the second ($F_1(1, 84) = 6.04$; $F_2(1, 48) = 5.21$) and third ($F_1(1, 84) = 8.52$; $F_2(1, 48) = 6.54$) word classes, but did not reach significance in the first word class ($F_1(1, 84) = 1.05$, $p = 0.31$; $F_2(1, 48) < 1$, n.s.).

7.3 *Results – summary*

The results of experiments 1 and 2 support the following conclusions.

(a) Homorganicity at the root's end is undesirable. Speakers consider roots with homorganic consonants to be less acceptable than controls. This conclusion agrees with the previous results of Frisch & Zawaydeh (2001) in Jordanian Arabic. Unlike these earlier studies, our experiments have systematically manipulated the position of C2C3 homorganic consonants within the word. Our Hebrew findings demonstrate the avoidance of homorganic root consonants across a variety of morphological templates, regardless of the position of the homorganic consonants in the word.

(b) The acceptability of identical consonants is inexplicable by similarity. Contrary to the prediction of the monotonic similarity account, roots with C2C3 identical consonants are more acceptable than roots with C2C3 homorganic, nonidentical consonants. In fact, these experiments each yielded higher ratings for root final identical consonants compared to homorganic nonidentical consonants, a trend that was significant in both experiments for the first word class. The greater stability of the effect for the first word class may be due either to the transparency of its morphological structure or to its frequency in the language. These factors may facilitate the extraction of the root, thereby increasing the salience of its internal structure. Sensitivity to root structure, however, emerged in other word classes as well. These results suggest that speakers encode the identity of root consonants, and their representation distinguishes between identical and homorganic consonants.

(c) Identity is not entirely desirable. Replicating previous results (Berent & Shimron 1997; Berent et al. 2001(a)), roots with identical consonants at their end were rated lower than controls. This last result is unexpected under the view that root final identity does not violate the OCP (McCarthy 1986). Although, as we explain below, the source of the relative unacceptability of *smm*-type roots remains uncertain, it is, nevertheless, patent that this finding is due to the identity of the C2C3 radicals, rather than to their similarity.

8. Discussion

Two experiments examined whether the acceptability of identical root consonants is an inverse, monotonic function of their perceived similarity. Although speakers are highly sensitive to similarity among root consonants, their ratings of identical consonants clash with the monotonic similarity explanation. We wish to make it perfectly clear that our experiments do not address the adequacy of the stochastic similarity model as an account for the restrictions on nonidentical root radicals. In fact, the similarity hypothesis may easily be modified to account for our findings by incorporating a separate set of constraints that specifically target identity. The merit of our conclusions, however, should not be judged by the repair they impose on alternative approaches but, instead, by their centrality to the domain of inquiry. The existence of grammatical constraints that specifically target full-segment identity is a question that is central to phonological theory. The phonological literature includes countless case studies implicating a distinction between full and partial identity in grammar. Our findings provide corroborative experimental evidence for such a distinction. These results suggest that speakers treat identity in a manner that is fundamentally different from partial similarity. To the extent that Semitic roots are subject to a stochastic constraint on perceived similarity, this constraint must be complemented by a separate set of grammatical principles or constraints concerning identity.

The role of identity in the grammar of Semitic has been the subject of some debate in the linguistic literature. McCarthy (1986) attributes the co-occurrence restrictions on identical root consonants to a constraint that specifically bans identity. Identity avoidance operates at the root domain and concerns adjacent radicals. The perceived adjacency of root radicals, in turn, is captured by segregating them from vowels and affixes. Each of these assumptions, however, has been met with some criticism. Pierrehumbert (1993) observed that identical root radicals are underrepresented at nonadjacent (C1C3) positions in Arabic roots, a finding that is inconsistent with the local definition of the OCP-total by McCarthy (see also Buckley 1997, for Tigrinya). Bat-El (1994) further rejected the OCP-total as an account for the distribution of adjacent identical consonants, suggesting that *samam*-type verbs are formed from monosyllabic nouns (e.g. *sam*) by means of reduplicative stem modification (see also Ussishkin 2000). More generally, Gafos (1998) criticized long-distance spreading on the grounds that it requires the segregation of consonants and vowels, in general, and, in the case of Semitic roots, the segregation of root consonants from vowels and affixes. Plane segregation incorrectly predicts unattested partial spreading of consonantal features across vowels (Gafos 1998). The proposals of Bat-El (1994), Gafos (1998) and

Ussishkin (2000) each capture the asymmetry in the distribution of identical consonants in Semitic verbs without appealing to either the root or identity avoidance. Additional evidence against the consonantal root challenges its role as a morphological constituent. Such observations suggest that Hebrew word formation is sensitive to the interdigitation of consonants and vowels. For instance, Bat-El (1994) observes that Hebrew verbs formed from nouns maintain the noun's clusters (e.g. ***pr**aklit*→***pr**iklet*). Likewise, the shape of roots in denominal Hebrew verbs is predictable from the vowel in the base (Ussishkin 1999, 2000).

These observations raise three questions with respect to the interpretation of our findings: (a) Do speakers of Semitic languages encode the identity of consonants relative to a morphological domain that coincides with the root? (b) How is this domain encoded? (c) Is identity banned by the grammar?

Our results clearly demonstrate that speakers can discriminate between novel words that differ solely in terms of the identity of root final consonants. Speakers are further sensitive to the presence of identical consonants regardless of their position in the word, and despite intermediate vowels and affixes (Berent & Shimron 1997; Berent et al. 2001(a)). For instance, *ssm*-type roots are unacceptable when the identical consonants are either word initial (e.g. *sisem*) or word medial (*histasem*), and when they are separated by various vowels (e.g. *i* or *a*) or an infix (e.g. -*ta*-; see Berent et al. 2002). The generalization of the restrictions on identical consonants to novel words is theoretically significant. Several existing accounts consider the asymmetry in the location of identical stem consonants in existing denominal verbs as an emerging property of anchoring the biconsonantal nominal base with the output (e.g. Bat-El 1994; Gafos 1998; Ussishkin 2000). These proposals successfully capture the distribution of identical root consonants without invoking grammatical constraints on identity. The asymmetry in the location of identical consonants for novel verbs, for which a biconsonantal base is not independently motivated, challenges this approach. It is, of course, perfectly possible that speakers infer a biconsonantal base for novel *smm*-type forms as well. The erasure of identity in novel bases calls for an explanation, however. Such an explanation may well require a grammatical constraint on full-segment identity. To capture the restriction on identical consonants across intermediate vowels and affixes, and irrespective of word position, the identity constraint must further operate within a morphological domain.

The nature of this morphological domain, however, is more difficult to ascertain. The perceived adjacency of identical C1C2 and C2C3 root radicals is handled naturally by the view of identity as constrained within the root domain, a phonological constituent that renders root consonants adjacent by segregating them from vowels. The segregation of the consonantal root in Semitic can further account for the prevalence of Arabic root-consonant metathesis errors in the output of a bilingual Arabic-French aphasic despite their rarity in his French output (Prunet, Béland & Idrissi 2000). As noted above, however, this approach predicts unattested effects of long-distance spreading (Gafos 1998). Furthermore, the restriction on consonant-identity is not entirely blind to intervening vowels. Previous experiments consistently demonstrate that root initial identical consonants are less acceptable when they are locally adjacent (e.g. *massimim*; the initial *ma-* and final *-im* are affixes), compared to when they are separated by vowels (e.g. *sisem*; Berent & Shimron 1997; Berent et al. 2001(a)). This finding suggests that the representation available to speakers specifies the interdigitation of consonants and vowels, an observation consistent with the proposals of Bat-El (1994) and Gafos (1998). To address these concerns, the identity of stem consonants must be captured without segregating the consonantal root from vowels.

Several existing accounts encode the identity of consonants in this fashion. Although these proposals are designed to capture distinct roles of identity in the grammar (i.e. identity avoidance vs. identity enforcement), they involve several strategies for encoding the identity between disjoint segments. One approach restricts identity between both adjacent and disjoint stem consonants by means of a family of identity-avoidance constraints (Everett & Berent 1998).[6] Although this approach captures the acceptability of identical root consonants in Hebrew (homorganicity is not specifically addressed), it does so by invoking numerous related constraints, each of which targets a distinct type of identity (the stem's C_1C_2 consonants, adjacent identical consonants and any identical stem consonants). An alternative approach assumes a single identity constraint that can be selectively applied to a pair of segments by virtue of their correspondence (Rose & Walker 2001). This approach enforces correspondence between (disjoint) similar consonants and favors full identity among these corresponding segments. Extending this general approach to Hebrew C_2C_3 consonants would correctly predict an advantage for full identity over similarity (provided that C_1C_2 identity is independently banned). A correspondence relation between identical stem consonants is also assumed by Bat-El (2002). This proposal accounts for speakers' ability to generalize the restrictions on long-distance identity to novel forms (i.e. without prior knowledge of the base). Bat-El (2002) suggests that the correspondence between identical stem consonants favors an interpretation such that one is a copy of the other, provided that the copied segment occupies the rightmost position in both the base and the stem. In summary, each of these three proposals offers a mechanism for restricting the co-occurrence of disjoint identical consonants without segregating them from vowels. Whether speakers encode the domain of identity-avoidance

[6] Although Everett & Berent's (1998) proposal invokes the root, there is nothing in their analysis that requires the representation of the root as a separate phonological constituent. Indeed, they specifically assume that identity formation is achieved by reduplication, rather than by long-distance spreading.

as the stem or the root – a phonological constituent that is segregated from vowels – is a question that falls beyond the scope of this discussion. Likewise, our discussion cannot address the role of the root in the morphology of Semitic, a question that is largely orthogonal to its role as a phonological constituent. Regardless of whether the co-occurrence restrictions on identical consonants operate over a root or a stem, however, speakers are clearly sensitive to long-distance identity and discriminate it from nonidentity.

Although the grammar of Hebrew appears to encode long-distance identity, it is uncertain whether it bans it. The experimental results demonstrate that roots with final identity are rated as relatively unacceptable compared to nonhomorganic controls. These results replicate our previous rating experiments (Berent & Shimron 1997; Berent et al. 2001(a)) indicating that root final identity is not entirely desirable. These findings also converge with the observations of Rose (2000), who documented long-distance avoidance of root final identical consonants (specifically, gutturals) in Tigrinya. Unlike the aversion to identical gutturals observed by Rose (2000), the unacceptability of root final identical consonants in Hebrew cannot be explained by homorganicity: if the unacceptability of *smm*-type roots was due to homorganicity, then this rejection should have been at least comparable to (in fact, larger than) the rejection of homorganic controls. Contrary to this prediction, *smm*-type roots are more acceptable than homorganic controls. The relative unacceptability of *smm*-type roots must then be due to their total identity.

Why are *smm*-type roots relatively unacceptable? One explanation attributes this finding to a grammatical constraint that bans long-distance identity (Everett & Berent 1998; Rose 2000). *Smm*-type forms are nevertheless attested because identity avoidance is dominated by higher-ranked constraints. For instance, Everett & Berent (1998) attribute the emergence of *simem* to the preference to fully align its biconsonantal base with the three consonant slots in the word pattern (i.e. MAXIO, see Everett & Berent 1998). The location of identity, in turn, is explained by the lower ranking of identity violation at the left edge of the root (Everett & Berent 1998). Conversely, according to Bat-El (1994, 2003), the emergence of long-distance identity in verbs such as *simem* is due to a constraint on the minimal word requiring bisyllabicity, whereas its location is explained by the leftward alignment of the base and the reduplicated output (but see Bat-El 2002 for a different approach). Despite their optimality with respect to higher ranked constraints, the identical consonants in *simem* violate an identity constraint (Everett & Berent 1998; Rose 2000; Bat-El 2003). The relative unacceptability of *smm*-type roots in our experiments may reflect such a violation of a grammatical constraint. However, the empirical evidence for identity violation in these forms is currently limited. To the extent that this experimental finding is not corroborated by converging linguistic evidence, an alternative explanation may be considered. This account attributes the relative unacceptability of *smm*-type roots to the rarity of XYY root types (relative to the XYZ type) in the lexicon, rather than to an active grammatical constraint.[7]

The question of whether long-distance identity is banned by the grammar must await further linguistic analysis. Our present goal, however, was to examine whether identity of root final consonants is represented and distinguished from homorganicity. The experimental evidence on this question is rather clear: the co-occurrence restrictions on identical consonants are inexplicable as a monotonic function of their similarity. Long-distance consonant identity, a formal relationship among segments, is thus irreducible to partial similarity at the level of individual features.

REFERENCES

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

Bat-El, O. (2002). *Hebrew reduplication: the interpretation of forms with identical consonants*. Paper presented at a colloquium, Department of Linguistics, University of Southern California.

Bat-El, O. (2003). Semitic verb structure within a universal perspective. In Shimron, J. (ed.), *Language processing and language acquisition in a root-based morphology*. Amsterdam: John Benjamins. 29–59.

Bender, M. L. & Fulass, H. (1978). *Amharic verb morphology, a generative approach*. East Lansing, MI: African Studies Center, Michigan State University.

Berent, I., Bibi, U. & Tzelgov, J. (2000). *The autonomous computation of linguistic structure in reading: evidence from the Stroop task*. A paper submitted to the 41st meeting of the Psychonomic Society.

Berent, I. & Shimron, J. (1997). The representation of Hebrew words: evidence from the Obligatory Contour Principle. *Cognition* **64**. 39–72.

Berent, I., Everett, D. & Shimron, J. (2001(a)). Do phonological representations specify formal variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology* **42**. 1–60.

Berent, I., Shimron, J. & Vaknin, V. (2001(b)). Phonological constraints on reading: evidence from the Obligatory Contour Principle. *Journal of Memory and Language* **44**. 644–665.

Berent, I., Marcus, G., Shimron, J. & Gafos, A. (2002). The scope of linguistic generalizations: evidence from Hebrew word formation. *Cognition* **83**. 113–139.

Bolozky, S. (1978). Some aspects of Modern Hebrew phonology. In Berman, R. (ed.), *Modern Hebrew structure*. Tel Aviv: Universities Publishing Projects. 11–67.

Buckley, E. (1997). Tigrinya root morphology and the OCP. *University of Pennsylvania Working Papers in Linguistics* **4.3**. 19–51.

[7] The sensitivity of our participants to type frequency could further be prompted by the explicit demand to provide acceptability ratings. Indeed, when well-formedness is assessed implicitly, there is no evidence that *smm*-type roots are unacceptable. For instance, consider again the lexical decision task (Berent et al. 2001(b)). Participants in this task are asked to determine whether or not a string of letters corresponds to an existing word by pressing one of two computer keys. Of particular interest are responses to novel words: the greater the ill-formedness of a novel word, the easier it should be for participants to classify it as a nonword. Specifically, if novel words with *smm*-type roots were ill-formed, then it should have been easier to classify them as nonwords, compared to matched controls with no identical consonants. Contrary to this expectation, nonword classification is significantly MORE difficult for *smm*-type roots relative to no-identity controls. Berent et al. (2001(b)) suggested that this difficulty reflects the perception of identity as a product of a grammatical operation, an attribute that renders novel stimuli more word-like. Clearly, however, the implicit measure of lexical decision yields no evidence for the unacceptability of *smm*-type roots.

Even-Shoshan, A. (1993). *Ha'milon ha'xadash* [The new dictionary]. Jerusalem: Kiryat Sefer.
Everett, D. L. & Berent, I. (1998). An experimental approach to the OCP: evidence for violable identity constraints in Hebrew roots. *Rutgers Optimality Archive* (ROA-235).
Frisch, S. (in press). Language processing and segmental OCP effects. In Hayes, B., Kirchner, R. & Steriade, D. (eds.), *Phonetic bases of markedness*. Cambridge: Cambridge University Press.
Frisch, S., Broe, M. & Pierrehumbert, J. (1997). *Similarity and phonotactics in Arabic*. Ms., University of Michigan, Ann Arbor.
Frisch, S. A. & Zawaydeh, B. A. (2001). The psychological reality of OCP-place in Arabic. *Language* **77**. 91–106.
Gafos, A. (1998). Eliminating long-distance consonantal spreading. *Natural Language & Linguistic Theory* **16**. 223–278.
Greenberg, J. (1950). The patterning of root morphemes in Semitic. *Word* **6**. 162–181.
McCarthy, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* **12**. 373–418.
McCarthy, J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry* **17**. 207–263.
McCarthy, J. (1994). The phonetics and phonology of Semitic pharyngeals. In Keating, P. (ed.), *Papers in laboratory phonology III*. Cambridge: Cambridge University Press. 191–283.
Pierrehumbert, J. (1993). Dissimilarity in Arabic verbal roots. In *Proceedings of NELS **23**, GLSA*. Amherst, MA: Department of Linguistics, University of Massachusetts. 367–381.
Prunet, J. F., Béland, R. & Idrissi, A. (2000). The mental representation of Semitic words. *Linguistic Inquiry* **31**. 609–648.
Rose, S. (2000). Rethinking geminates, long-distance geminates and the OCP. *Linguistic Inquiry* **31**. 85–122.
Rose, S. & Walker, R. (2001). *A typology of consonant agreement as correspondence*. Ms., University of California at San Diego & University of Southern California.
Sandler, W. (1994). *Suboral articulation*. Ms., Haifa University.
Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence. *Phonology* **16**. 401–442.
Ussishkin, A. (2000). *The emergence of fixed prosody*. Ph.D. dissertation, University of California at Santa Cruz.

Authors' address: Department of Psychology, Florida Atlantic University,
777 Glades Road, Boca Raton, FL 33431-0991, U.S.A.
E-mail: Iberent@fau.edu