ROBERT D. HOBERMAN

LOCAL AND LONG-DISTANCE SPREADING IN SEMITIC MORPHOLOGY

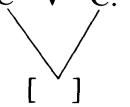
1. INTRODUCTION

In a Semitic language, as McCarthy (1981) has shown, consonantal roots, vocalic melodies, and CV templates are separate morphemes lying on separate autosegmental tiers, and are autosegmentally linked by universal conventions and language-particular rules. Most work in autosegmental phonology and morphology has assumed that the automatic linking of unassociated elements applies from left to right. When there are more C slots in a template than consonant phones in a root, it would normally be the last C slot that remains unassociated. Therefore if spreading is to apply, it would normally be expected to involve the last root consonant.

In fact, this is not the most widespread pattern in Semitic morphology. It will be shown here that left-to-right linkage and spreading *normally fail* when they would create transvocalic spreading; this is so cross-linguistically and without morphological stipulation. There is a general preference for LOCAL SPREADING – the multiple association of a consonant with two C slots that are adjacent, forming a geminate cluster – rather than LONG-DISTANCE SPREADING – multiple association across a V slot. This is so regardless of whether the first, middle, or last root consonant is the one to spread. Morphological and phonological processes frequently serve to avoid long-distance spreading, while many times creating local spreading. Invariably, when long-distance spreading does exist, the same CV template occurs also without spreading when the number of consonant phones is sufficient. These substantive properties imply the following markedness relationship:

(1) LOCAL SPREADING MARKEDNESS RELATIONSHIP (LSMR)
The configuration C C is more highly valued

than the configuration C V C.



The phenomena described by the LSMR cannot be derived from the

Natural Language and Linguistic Theory **6** (1988) 541–549.

well-known fact that geminates are not split by rules of epenthesis, or from the formal principles which have been shown by Hayes (1986), McCarthy (1986), and Schein and Steriade (1986) to underlie that fact. In the following sections it will be shown that local spreading is a favored output target in various morphological structures in four Semitic languages – Syriac, classical Arabic, colloquial Arabic, and biblical Hebrew – and that the LSMR motivates diachronic changes as well. Limitations of space make it impossible to present here a complete analysis of each of the languages from which arguments are brought. The details, including a full specification of the rules that are involved, an examination of all form classes, and the treatment of morphologically and lexically specified exceptions, are given in Hoberman (1987).[1]

2. Syriac

In Syriac, as in other Semitic languages, all productive stem templates contain three or more C slots. Sample verb forms containing the biconsonantal root **pk** 'smash' are listed in (2) and (3), juxtaposed with the corresponding forms from the triconsonantal root **ktb** 'write':[2]

(2)	pekk-et	ketb-et	Class I Perfect 1SG
	et-pakk-at	et-katb-at	Class I-t Perfect 3SG FEM
	t-et-pakk-ūn	t-et-katb-ūn	Class I-t Imperfect 2PL MASC
	pākk-ā	kātb-ā	Class I Act. Participle SG FEM
(3)	te-ppok	te-ktob	Class I Imperfect 2SG MASC
	te-ppok-ūn	te-ktob-ūn	Class I Imperfect 2PL MASC
	appek	akteb	Class IV Perfect 3SG MASC
	ett-appak	ett-aktab	Class IV-t Perfect 3SG MASC
	ett-appak-at	ett-aktab-at	Class IV-t Perfect 3SG FEM

Each of the categories listed here has the same CV structure for both biconsonantal roots (**pk**) and triconsonantal roots (**ktb**). In the former

[1] I am grateful to Mark Aronoff, Ellen Broselow, the *NLLT* Associate Editor Michael Kenstowicz, and three anonymous referees for very helpful critiques of earlier versions of this paper.

[2] In Syriac and Biblical Hebrew, postvocalic nongeminate, nonemphatic stops are realized as fricatives. In Syriac, following the application of this rule of Spirantization, lax vowels in open syllables delete: **te-ktob-ūn** [te-xtv-ūn]. Three of the six derivational classes (binyanim) of the Syriac verb are labeled here with the Roman numerals of the cognate Arabic classes: I, II, IV; the remaining three are the corresponding intransitives prefixed with **et(t)**-: I-t, II-t, IV-t.

case, one of the root consonants is geminated. In group (2), it is the second consonant **k** of the biconsonantal root that is geminate, parallel to the adjacent **tb** in the triconsonantal verb; corresponding to **ketbat** we have biconsonantal **pekkat**, not ***pepkat**. The generally accepted one-to-one, left-to-right association conventions of Pulleyblank (1986), with the addition of a language-particular rule of spreading, correctly derive these forms. In (3), however, it is the first root consonant **p** that is geminated, parallel to the adjacent consonants **kt** in the triconsonantal root. Thus the biconsonantal form corresponding to **tektob** is **teppok**, not ***tepkok**. Left-to-right linking and spreading would give the incorrect forms ***tepkok, *apkek, *ettapkak**, etc.

In both sets (2) and (3) the correct forms exhibit local spreading, **-pp-** or **-kk-**, rather than long-distance spreading, ***-kVk-** or ***-pVp-**. We could postulate a special rule which would have the effect of converting forms like ***tepkok** into the correct **teppok**. However, a more unified analysis of Syriac morphology can be developed if we assume that the first step in the linking of root phones with C-slots is to anchor both the initial and the final stem consonants. Prince (1987, p. 504) proposed a rule that associates a final root consonant to the final C slot in the CV template. Yip (1987) has generalized this notion, arguing that the universal, unmarked mode of association is not from left to right but from the edges inward. (Both of these proposals were based on Arabic data.) In the cited Syriac forms, edge-in association leaves a medial C slot unassociated: **peCk-at, te-pCok;** in both cases, a rule of Gemination (local spreading) applies, giving **pekkat, teppok.**

In morphological classes which have four C slots, even triconsonantal roots undergo spreading. Thus the Class II Perfect is **katteb,** not ***katbeb** as left-to-right association would have it. The correct form is derived automatically by the analysis proposed for biconsonantal roots. Edge-in association anchors the first and last stem consonants, and the medial geminate is produced by the rule of Gemination. Class II forms with biconsonantal roots (**pakkek**) are derived by the same mechanism, with the additional stipulation that it is the final consonant, rather than the initial one, that spreads (***pappek**).[3]

The same patterns of association are found not only in the verb system but also in nouns and adjectives. Typical examples are **rabbā** 'great', **ma-ʕʕālē** 'entrance', **rukkābā** 'combination', **sebbel-tā** 'ladder', **zakkāy** 'pure', **dabbābā** 'a fly', **rakkīk** 'soft'. For every association pattern found

[3] The same stipulation is involved in the derivation of forms which contain long-distance spreading, such as **et-pakek** Class I-t Perfect 3SG MASC.

in verb morphology, analogous patterns exist in nominals, with their much larger variety of CV templates and vowel formatives, making unworkable any notion that the verb patterns are due to morphologically-specified rules.

In some Syriac forms, long-distance spreading is avoided without producing local spreading. An example is the Class I Active Participle, for which a triconsonantal verb has the form **kāteb**, but a biconsonantal verb has **pāek**, not ***pākek** (as in most other Aramaic languages) or ***pāpek**. Syriac has a rule that deletes the middle, unassociated C slot in this and related forms. Although the consonant deletes in the masculine singular **pāek**, it does not in suffixed forms, FEM SG **pākkā**, MASC PL **pākkīn**, FEM PL **pākkān**. These have the same underlying CVCVC template as **pāek**, but the second stem vowel **e** is deleted by a rule of syncope; as a result, the conditions exist for the rule of Gemination to apply, bleeding Consonant Deletion.

3. Classical Arabic

In Classical Arabic, unlike Syriac, verbs with biconsonantal roots have forms with surface CV structures which differ from those of triconsonantal verbs. They have geminate consonants where the triconsonantal verb has an additional syllable. Compare the representative forms of **ktb** 'write' and **rd** 'return' in (4):

(4)				
Perfective	3 MASC SG	katab-a	radd-a
	1 PL	katab-naa	radad-naa
Imperfective	3 MASC SG	ya-ktub-u	ya-rudd-u
	3 FEM PL	ya-ktub-na	ya-rdud-na

Before vowel-initial suffixes, the **rd** verb has a geminate consonant **d**, e.g. **radda, yaruddu**. These forms have been treated by McCarthy (1981, p. 398; 1986, pp. 211, 246–248) as derived from templates identical to those of triconsonantal verbs via a rule which has the effect of metathesis in some cases (**yardudu → yaruddu**) and syncope in others (**radada → radda**). No change takes place if a consonant-initial suffix follows, as in **radadnaa**, **yardudna**, or there is no suffix, as in the jussive **yardud**. The rule is complex in form, performing different operations under different conditions; it need not be quoted here because its workings are not at issue. What is noteworthy is that in spite of this complexity, every application of the Metathesis/Syncope rule replaces an underlying long-distance spreading structure with local spreading. Thus the existence of

this rule is motivated by the marked status of long-distance spreading configurations.

4. Modern Colloquial Arabic

The Classical Arabic Metathesis/Syncope rule discussed in the previous section has the effect of creating stem alternations, **radd-** ~ **radad-** and **-rudd-** ~ **-rdud-,** and thus of diminishing the uniformity of the stem shapes in the paradigm. Many, perhaps all, modern vernacular Arabic dialects have leveled the paradigm. To do this, they had the alternatives of generalizing either of the two stem shapes in each tense/aspect paradigm. In every instance they generalized the stem shape with local spreading, not the one with long-distance spreading, thereby going further than Classical Arabic in complying with the LSMR. Compare Classical **radadnaa** 'we returned' with Cairene **raddeena**. The change from the Classical to modern colloquial forms consists of a generalization of the rule of Metathesis/Syncope and the addition of a rule which inserts the vowel **ee** between a biconsonantal stem like **radd-** and a consonant-initial subject suffix: **radd-ee-na.**

In the absence of the LSMR, some dialects should be expected to have taken the route of generalizing the long-distance spreading forms, producing 3 MASC SG ***radad**, 3 FEM SG ***radadat,** 3 PL ***radadu**, Imperfective ***yardud**, but I know of no dialect which has generalized such forms, or which has failed to generalize the stem shapes **radd-, (ya-)rudd**.

5. Biblical Hebrew

In Biblical Hebrew, verbs with biconsonantal roots have two sets of stem shapes, one following the syllable structure of triconsonantal verbs, the other deviating from it. The choice of one pattern or the other is in part free variation, in part lexically specified, and in part determined by a variety of syntactic and semantic factors. For many verbs, both patterns are attested in identical or nearly identical inflectional categories, for example **zāmam-tī** and **zamm-ō-tī** 'I intended'; **wə-ha-ħtat-tī** 'I will break' and **ha-ħitt-ō-tā** 'you broke'; both forms occur in the same sentence in **sabb-ū-nī gam səbāb-ū-nī** 'they surrounded me, indeed they surrounded me'. Forms of both types occur in all the relevant morphological categories: suffixed and unsuffixed forms of all tenses, participles, and infinitives, in verbs of all the derivational classes which do not preclude them for phonological reasons.

Whatever may be the historical origin of this formal duality, biradical verbs have, synchronically, two alternative CV-templates. The first, which contains long-distance spreading (e.g. **zāmamtī**), is identical to the template for verbs with triconsonantal roots (**kātabtī**), and thus exists in the grammar independently. The other (**zammōtī**) is special to biconsonantal verbs, and is derived, according to an analysis by McCarthy (1986, p. 237), by a rule of 'Geminate Verb Deletion' which deletes a vowel between colinked consonants (and is part of, or a replacement for, the rule of Metathesis/Syncope). Furthermore, there is a rule that adds **ō** in these forms just as **ee** is added in colloquial Arabic dialects. The cost to the grammar of these rules and consequent paradigm heterogeneity is apparently offset by the fact that they generate forms with local rather than long-distance spreading.

6. Class II Verbs in Arabic

An Arabic verb of Classes II or V, like the corresponding Syriac classes mentioned above, always has a geminate middle consonant: **kattab, takattab**. Left-to-right linkage and spreading would yield the incorrect result ***katbab**. To derive the correct form McCarthy (1981, p. 392) postulated a morphologically restricted rule, 'Second, Fifth Binyanim Erasure' (II/V Erasure), to erase the association of the third C slot, producing **katCab** with one C slot in the prosodic template unassociated. Such a C slot is associated with the same melody element as the C to its immediate left, producing the correct form **kattab**. However, several kinds of synchronic and diachronic evidence show that this approach is incorrect, and that the gemination of the middle consonant in Classes II and V is not morphologically stipulated at all.

A. Classes II and V are the only triconsonantal verb classes with the base template CVCCVC (excluding prefixes). There are numerous nouns and adjectives, however, with the templates CVCCVC and CVCCVVC, and in virtually all of them it is the middle root consonant, not the final one, which spreads. Typical examples include **qullab** 'changeable', **tujjaar** 'merchants' (pl. of **taajir**), **xabbaaz** 'baker', **baṭṭiix** 'watermelon (collective)', **ṣiddiiq** 'righteous'. There is no simple morphological characterization of all these types. Thus the erasure rule restricted to verb classes II and V is not adequate; rather all these nouns and adjectives as well as verbs of classes II and V behave in autosegmental association as they do for phonological, not morphological, reasons.

Class II may be compared with Class IX (e.g. **hmarar-** 'be, become red'), which, unlike II, would be produced directly by the standard

left-to-right association convention. On the present analysis, with edge-in association, the derivational complexity of the two classes is equal.[4] In fact, it can easily be shown that Class II, with local spreading, is formally and semantically less marked than Class IX, which has long-distance spreading, suggesting that both the LSMR and edge-in association are valid.

B. The range of functions of the Class II derivational pattern is very broad: Class II verbs are basic, causative, or intensive; they may be denominative; though usually transitive, some are intransitive. Class II is the second broadest in function and probably the second most numerous verb class in Arabic. Class IX verbs on the other hand are extremely restricted. Almost without exception, Class IX verbs are intransitives expressing colors and bodily defects: 'to be or become red', 'to be or become crooked', and so on; their number is naturally limited. Class II verbs appear in virtually every paragraph of Arabic; whole books might not contain a single Class IX verb. Thus it is Class IX, not Class II, that is semantically and syntactically marked.

C. Diachronic evidence supports this conclusion. Class II verbs exist in every dialect of colloquial Arabic, and indeed in some dialects, such as Syrian, this class has expanded its range of functions, supplanting Class IV for causatives. Class IX, on the other hand, has been lost in at least one dialect group, Moroccan. In some Moroccan dialects (Meknes), Class IX verbs have been reassigned to Class V, which is like Class II in having a geminate *middle* root consonant, so that for example classical **iḥmarra** 'it became red', **iḥmararnaa** 'we became red', have become **tḥemmeṛ**, **tḥemmeṛna** respectively, and long-distance spreading has been replaced by local spreading. In other dialects, however, such as those described by Harrell (1962), Class IX survives as a semantic and derivational class but its form has changed, eliminating multiple association entirely: corresponding to classical **iḥmarra**, **iḥmararnaa**, these dialects have **ḥmaṛ**, **ḥmaṛna**. This shows that the extinction of the Classical Arabic type of Class IX is due not to its semantic narrowness but to its phonologically marked character.

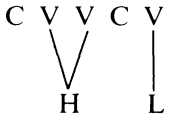
D. Furthermore, Class II exists in all branches of the Semitic family, both ancient and modern, and so can be ascribed with certainty to proto-Semitic. Thus it has survived for at least five or six thousand years. Either the II/V Erasure rule of McCarthy (1981) or the edge-in asso-

[4] The set of rules that embodies this analysis correctly derives all the Arabic verb classes except for Classes XII and Q-IV, which require morphological stipulation of the correct linkage configurations.

ciation convention of Yip (1987) with the rule of Gemination has persisted throughout that period. II/V Erasure is a language-particular, morphologically conditioned rule; hence it should be possible for some language to lose the rule, with the effect that Class II verbs like **darras** would be reshaped as ***darsas**. No Semitic language has made this change. On the other hand, not only is the form of Class II **darras** motivated by the LSMR, but edge-in association is proposed not as a rule at all but rather as a universal convention. Therefore the longevity of Class II is much better attributed to the latter analysis than to the erasure rule. This in turn supports the LSMR.

7. Beyond Semitic

The languages examined here are all genetically related and typologically similar, and it remains to be seen whether the Local Spreading Markedness Relationship can be shown to apply in dissimilar languages. There are few, if any, languages outside the Semitic family that have consonants and vowels on separate morpheme tiers, making long-distance spreading of consonants possible. However, it may be appropriate to extend the LSMR not only to vowels, but to all autosegmental association phenomena. The LSMR, if generalized to apply to tone structures, would predict that if a language has tonal melodies and long vowels, then level tones, as in (5a), should be preferred over contour tones (5b):

(5)a. 

b. 

If this is true, it would strongly support the LSMR.

References

Harrell, William Slade: 1962, *A Short Reference Grammar of Moroccan Arabic*, Georgetown University Press, Washington, D.C.

Hayes, Bruce: 1986, 'Inalterability in CV Phonology', *Language* **62**, 321–351.

Hoberman, Robert D.: 1987, 'Local Spreading', unpublished ms., SUNY at Stony Brook.

McCarthy, John J.: 1981, 'A Prosodic Theory of Nonconcatenative Morphology', *Linguistic Inquiry* **12**, 373–418.

———: 1986, 'OCP Effects: Gemination and Antigemination', *Linguistic Inquiry* **17**, 207–263.

Prince, Alan S.: 1987, 'Planes and Copying', *Linguistic Inquiry* **18**, 491–509.

Pulleyblank, Douglas: 1986, *Tone in Lexical Phonology*, Reidel, Dordrecht.

Schein, Barry, and Donca Steriade: 1986, 'On Geminates', *Linguistic Inquiry* **17**, 691–744.

Yip, Moira: 1987, 'Edge-In Association', *Proceedings of NELS 18*, University of Massachusetts at Amherst; revised version 'Template Morphology and the Direction of Association', *NLLT*, this issue, 551–577.

Received 6 June 1986
Revised 20 January 1988

Program in Judaic Studies
State University of New York at Stony Brook
Stony Brook, NY 11794–3355
U.S.A.