University of Pennsylvania Working Papers in Linguistics

Volume 4 | Issue 3 Article 3

1-1-1997

Tigrinya Root Consonants and the OCP

Eugene Buckley

Tigrinya Root Consonants and the OCP

Tigrinya Root Consonants and the OCP

Eugene Buckley

The Obligatory Contour Principle (OCP), which expresses a prohibition on adjacent identical elements, has played an important role in the development of phonological theory. Originally proposed in the domain of tone, it has received some of its most striking support from the nature of consonantal roots in Semitic languages. For example, in Classical Arabic there are no roots which consist of identical adjacent consonants *(*qql),* a fact which is attributed to the OCP. In its standard formal conception, the OCP is an absolute principle: it judges between identity or lack of identity, permitting only the latter cases. In fact, however, the Semitic data support a more gradient interpretation: while identical adjacent consonants are prohibited, nonadjacent identical consonants are disfavored but attested; so are adjacent nonidentical (but homorganic) consonants.

In this paper I survey evidence from the Ethio-Semitic language Tigrinya which supports these conclusions. I begin in §1 by describing the root-and-pattern morphology of Semitic languages, together with the consonantal roots which constitute a central element of the system. In §2 I survey traditional observations regarding restrictions on the cooccurrence of similar consonants within a root, and how the OCP has been applied to explain them. In §3 I describe the corpus used in the study, outline the methodology, and discuss the major results. In §4 I show how Pierrehumbert's (1993) similarity model accounts for the observed patterns. In §5 I consider an additional hypothesis regarding the role of root length in cooccurrences. A brief conclusion is given in §6.*

1. Roots and Templates

§§1–2 are a review of Semitic templatic morphology and the role of the OCP in phonological theory, respectively, which also makes my assumptions fully

* Results in this paper were presented at the Penn Linguistics Colloquium. University of Pennsylvania, February 20-21, 1993; and the Annual Conference on African Linguistics, Ohio State University, July 23-25, 1993. I would like to thank Mark Liberman for crucial programming assistance, and Stefan Frisch, John McCarthy, and Janet Pierrehumbert for helpful comments. All errors are, of course, my own.

explicit. Readers who are already familiar with these topics can skip ahead to §3 where the core of the Tigrinya pattern is presented.

A fundamental characteristic of Semitic morphology is the use of consonantal roots and syllabic patterns, or **templates**, in various combinations. I use data from Tigrinya to illustrate, but the phenomena are typical for the family as a whole. For example, the following words illustrate the realization of the root /sbr/ 'break (tr.)' in a range of templates (Leslau 1941, Berhane 1991); in many cases there are also affixes.

(1)	a.	sʌbʌr-ku	'I broke (it)'
	b.	yɨ-ssʌbʌr	'may it be broken'
	c.	tɨ-sʌbbɨr	'she breaks (it)'
	d.	nɨ-sʌbr-o	'we break it'
	e.	sɨbʌr	'break (it)! (m.sg.)'
	f.	yɨ-sbʌr	'may he break (it)'

The data in (1) reflect various inflectional categories for a single verb, whose exponence is not only in prefixes and suffixes but also in the shape of the **stem**: that is, the combination of the consonantal root with the template, which in Tigrinya includes information about the distribution of vowels, consonants, and consonant gemination. Thus the template for the perfect in (1a) can be represented abstractly as *CʌCʌC*. This template happens to occur with an obligatory set of suffixes which indicate the person, number, and gender of the subject, as in (2).

(2)		*singular*	*plural*
1st person		sʌbʌr-ku	sʌbʌr-na
2nd person	*masc.*	sʌbʌr-ka	sʌbʌr-kum
	fem.	sʌbʌr-ki	sʌbʌr-kɨn
3rd person	*masc.*	sʌbʌr-ʌ	sʌbʌr-u
	fem.	sʌbʌr-ʌt	sʌbʌr-a

Similarly, each of the other templates in (1) occurs with a particular set of affixes. My interest in this paper, however, is on the form of the stem, independent of the affixes. I will therefore generally omit affixes in illustrating stems, indicating the incomplete status of a stem by a hyphen.

In addition to the inflectional templates in (1), templates are also used for derivational purposes. For example, a root which can be used in verb patterns can often also be used as a noun in a different templatic pattern.

The pairs below, from Bassano (1918), illustrate the roots /dnk̋/, /drˀ/, and /ˀgs/.

(3) a. dʌnnʌk̋- ‘be astonished’
b. dɨnk̋i ‘wonder, surprise’

(4) a. dʌrrʌˀ- ‘patch up, reinforce with patches’
b. dɨrˀ-ɨto ‘patchwork quilt’

(5) a. ˀɨggus ‘patient’
b. tɨ-ˀgɨs-ti ‘patience’

In such cases, the consistent correlation between consonants and semantics motivates the positing of the consonantal root as an independent morpheme. It is the nature of this consonantal root which is the focus of this paper.

1.1. One-to-One Correspondences

All examples given so far involve roots with three consonants, or **triliterals**; this is by far the most common root type throughout Semitic, including Tigrinya. Additional examples are given below, in the perfect stem. In (6) are Type A verbs, the default class, while in (7) are verbs from Type C, a small class characterized by conjugations with the vowel [a] between the first two consonants (Leslau 1941: 95, Berhane 1991: 168ff).

(6) a. gʌrʌf- ‘whip’
b. wʌlʌd- ‘give birth’
c. lʌwʌs- ‘mix’
d. bʌlˀ- ‘eat’

(7) a. barʌk- ‘bless’
b. laš̋ʌy- ‘shave’
c. safʌg- ‘hesitate’

While this root type is the most common, both shorter and longer roots also exist. Following are the perfect stems of some **quadriliteral** roots, containing four consonants.

(8) a. mʌskʌr- ‘witness’
b. gʌlbʌt̓- ‘turn’
c. šʌmgʌl- ‘mediate’

True **quinquiliterals**, i.e. roots with five underlying consonants, are exceedingly rare; Bassano (1918) gives only one root which is treated here as quinquiliteral. See the next section for discussion.

1.2. One-to-Many Correspondences

For the basic perfect stems examined so far, the number of root consonants is the same as the number of consonant slots in the template. For example, the common Type A verb exemplified by *sʌbʌr-* ‘break’ has a triliteral root with a three-consonant template. Similarly, the verbs in (8) have quadriliteral roots and four-consonant templates. But in many cases the number of root consonants is smaller than the number of consonant slots which they need to fill. For example, the common Type B verb is characterized by a triliteral root and a four-consonant template. To fill the extra consonant slot, this verb type has gemination of the medial consonant.

(9) a. bʌddʌl- ‘hurt’
b. rʌssɨʔ- ‘forget’
c. wʌddʌs- ‘praise’

Less common, but reflecting the same combination of a triliteral and four slots, are verbs with spreading of the final consonant to fill the remaining slot.

(10) a. dʌrsʌs- ‘heal’
b. saħbʌb- ‘become exhausted’
c. lʌmt̓ʌt̓- ‘beat soundly’

The contrast between the verbs in (9) and (10) is illustrated below, using the templatic formalism introduced by McCarthy (1981).

(11) a.
```
C ʌ C C ʌ C     =     bʌddʌl
|    \/    |
b    d     l
```

```
b.  C ʌ C C ʌ C      =       dʌrsʌs
    |     |  \  /
    d     r   s
```

There are various theories about how to derive the difference between the two verb types and the way they are realized in the template (see Buckley 1990 for Tigrinya); the important point here is that both verbs have three root consonants. That is, at the consonant level of representation, they are of exactly the same triliteral type, and are equivalent for the purposes of this investigation.

Except for three irregular verbs—*hab-* 'give', *ħaz-* 'hold', *bʌl-* 'say'—there are no templates with just two slots for consonants.[1] This means that when a **biliteral** root, with just two consonants, is associated to a template, at least one of the consonants has to serve double duty, as in (11). A very common pattern in Semitic is the repetition of the second consonant. This can occur with a three-slot template, as in (12).

(12)	a.	sʌdʌd-	'send'
	b.	kʌbʌb-	'surround'
	c.	mʌzʌz-	'draw sword'
	d.	nazʌz-	'forgive'

Similar (though less common) repetition can occur when a biliteral root is combined with a four-slot template, where the three final slots all instantiate one consonant.

(13)	a.	k̓ʌnnʌn-	'pour off liquid'
	b.	gʌddʌd-	'be important to (someone)'
	c.	sʌbbʌb-	'become moldy (of bread)'
	d.	mʌssʌs-	'try, begin'

These are generally considered a subclass of Type B, since the medial consonant is geminated (Leslau 1941: 109). More typical when a biliteral associates to four slots is that the set of two consonants is reduplicated, creating what is often treated as a subclass of quadriliterals (Leslau 1941: 96).

[1] Even the irregular verbs can be shown to derive from the triliteral roots /whb/, /tħz/, and /bhl/ (Leslau 1941: 122f). Other apparent two-consonant stems such as *kʌd-* 'go' and *mot-* 'die' involve deletion or coalescence of a glide (the roots are /kyd/, /mwt/).

(14) a. mʌrmʌr- 'examine'
b. lʌflʌf- 'chatter uselessly'
c. lʌmlʌm- 'be soft'

What all the roots from (12) to (14) have in common is that they consist underlyingly of just two consonants. As a result, they are, like the roots in (11), equivalent to each other for the purposes of the root-consonant cooccurrences investigated below.

Finally, there are certain verbs which involve a triliteral root linked to a template with five consonant slots. These entail reduplication of the final two consonants, similar to (14). Every five-slot template in Tigrinya includes a prefix such as causative *ˀa-* or passive-reflexive *tʌ-*.

(15) a. ˀa-ˀrʌmrʌm- 'resent'
b. ˀa-htʌftʌf- 'talk in sleep, delirium'
c. ˀa-zfʌrfʌr- 'hang heavy (with fruit)'
d. tʌ-k̓bʌzbʌz- 'look around curiously'

Abstracting away from the reduplicated part of the stem, these verbs are derived from triliterals such as /ˀrm/, and are assimilable to the large class of triliteral roots more typically realized as in (6). Only one root, identified by Bassano as belonging to the Hamasen dialect, is treated in this analysis as a true quinquiliteral (cf. also (37) below). It is also unique among five-slot templates in that it takes no prefix.

(16) gʌrʌngʌr- 'start to form pod (of legume)'

While this root might need to be treated synchronically as /grngr/, it is presumably the result of an irregular reduplication of /grn/. Due to its multiply exceptional nature, it is not included in the statistical analysis here.

2. The Obligatory Contour Principle

2.1. Identical Consonants

In (12) above we saw that when a biliteral root such as /sd/ associates with a three-slot template, it is the second consonant which occupies two slots (17a), rather than the first (17b). That is, while stems of the type *sʌdʌd-* are common, those of the form **sʌsʌd-* are absent.

(17) a. C ʌ C ʌ C = sʌdʌd

s d

b. C ʌ C ʌ C = *sʌsʌd

s d

This pattern is well known for Semitic and to some degree Afro-Asiatic (Greenberg 1950, Bender 1978), and requires a general solution. The standard explanation in modern phonological theory (McCarthy 1986) has two parts. First, as assumed tacitly in the discussion up to this point, roots cannot have two identical consonants in a row. That is, both */sdd/ and */ssd/ are prohibited as underlying roots. This effect is attributed to a general phonological notion motivated by tone and many other areas (Leben 1973, Goldsmith 1976, McCarthy 1986).

(18) *Obligatory Contour Principle (OCP)*

Adjacent identical elements are prohibited.

The OCP requires that a stem like *sʌdʌd-* be derived from the root /sd/, since */sdd/ is a violation of the principle, just as */ssd/ is a violation.

The responsibility for ensuring that it is the **second** consonant that spreads to two consonant slots then falls to the association algorithm, i.e. the rules that link up the root consonants to the template. The essential element of the explanation is that the consonants link up to the template slots one at a time, from left to right.

(19) *Left-to-right association*

C ʌ C ʌ C

s d

The second step is spreading of the rightmost consonant to the remaining slot, yielding the correct output *sʌdʌd-*.

(20) *Rightward spreading*

```
C ʌ C ʌ C
|   | ,-'
s   d
```

If association is always left to right, the unattested form **sʌsʌd-* will not be derived, and the asymmetric pattern is correctly captured.[2]

2.2. Identical Place Features

In the preceding section we saw how the OCP operates at the segmental level (root tier) to prohibit adjacent identical consonants. This can be termed the ‘total OCP’ (Pierrehumbert 1993), since it refers to adjacent consonants that are identical in all their features. But the OCP is also important in explaining another widespread generalization regarding Semitic roots. As observed by Greenberg (1950) and others, not only are adjacent identical consonants prohibited within a root, but even nonidentical consonants of the same place of articulation, and even those which are nonadjacent, are strongly disfavored within a root. In other words, a ‘place OCP’ appears also to operate on individual place features to prohibit same-place consonants anywhere in the same root.

First it is necessary to clarify the category ‘place of articulation’. For the purposes of this generalization, the following classes of consonants are relevant; the consonants which make up each class are given for Tigrinya, but the same basic classes hold across the languages.

(21)

gutturals	h ʔ ħ ʕ
velars	k g k̓ kʷ gʷ k’ʷ
coronal obstruents	s z s̓ t d t̓
coronal sonorants	r n l
labials	f p b p̓ m

Thus while a root like */kbb/ is prohibited by the total OCP, there is also a strong ‘place OCP’ effect which disfavors roots such as /kbm/—adjacent and homorganic but nonidentical—and /bkb/—identical but nonadjecent. There

[2] See McCarthy (1981), Yip (1988), and Buckley (1990) for various analyses of other stem types such as *bʌddʌl-*.

is, further, a weaker effect disfavoring nonidentical, nonadjacent consonants, e.g. /bkm/. (Note: throughout this paper, “nonidentical” is used to refer to consonants which are homorganic but not identical in all other features.)

Since homorganicity is crucially involved in these generalizations, a proposed explanation of the cause is that, as suggested, the OCP operates not only on the consonant as a whole, but also on individual place features. The analysis is couched in a feature geometry framework which has privative articulator nodes which occur on different tiers of the representation (Sagey 1986, Mester 1986, McCarthy 1988, Yip 1989). The four articulator nodes Pharyngeal, Dorsal, Coronal, and Labial encode the four basic place classes in (21). Adjacency on an articulator tier is unimpeded by intervention of a different-place consonant, permitting consonants which are not strictly adjacent to interact in a local fashion at the level of articulator. In (22), notice that the consonants /f/ and /b/ both bear a Labial node, and that these nodes are adjacent to each other on their tier regardless of whether a consonant such as /g/ intervenes, since /g/ has no Labial node.

```
(22)    g    f    b              f    g    b
        |    |    |              |    |    |
        |   Lab  Lab            Lab   |   Lab
        |                             |
       Dor                           Dor
```

This approach has the weakness that it is categorical: it predicts that /gfb/ and /fgb/ should be equally disfavored, since the Labial nodes that constitute the OCP violation are equally adjacent in each case. However, as Pierrehumbert (1993) shows for Arabic, the degree to which coocurrence of homorganic consonants is disfavored correlates with their proximity in the root: /gfb/ is worse than /fgb/. Further, the degree of disfavoring is proportional to the relative identity of the homorganic consonants in features beyond those of place of articulation. For example, as we shall see, /gtd/ is worse than /gtz/, since the coronal obstruents in the first root share the feature [–continuant], while those in the second root do not.

In the next section I replicate Pierrehumbert’s Arabic results using data from Tigrinya, following her fundamental approach and assumptions. Differences of approach will be pointed out as they arise.

3. Tigrinya Root-Consonant Cooccurrences

3.1. The Corpus

The data in this section are drawn from a corpus of all the verb roots found in Bassano (1918), the most complete dictionary of Tigrinya available at present. The consonantal roots were entered into a computer database for ease of seaching and analysis. For the purposes of this investigation, the template to which a root associates has been ignored (cf. §1.2). For example, while the following verbs have various templates, as well as different ways of associating to the template (spreading vs. reduplication), they are all treated here as underlyingly biliteral /zl/ or /mz/.

(23)	a.	zʌlʌl-	'jump (over an obstacle)'
	b.	ˀa-zlʌl-	'wear no pants, just a robe' *(caus.)*
	c.	zʌllʌl-	'dilute with too much water'
	d.	zʌlzʌl-	'cut meat in strips for drying'

(24)	a.	mʌzʌz-	'draw sword'
	b.	mʌzzʌz-	'belong, be fitting'
	c.	mʌzmʌz-	'be very long (of incisors); burn'

Similarly, the following verbs are treated as triliteral /sbl/ and /ħnk̉/ despite the templatic differences.

(25)	a.	sʌbbʌl-	'sprout'
	b.	sʌblʌl-	'load onto pack animal'

(26)	a.	ħanʌk̉-	'strangle'
	b.	ħank̉ʌk̉-	'be spoiled, finicky (of child)'

It may be that the template to which a root associates also affects cooccurrence patterns—for example, Lightner (1973: 58f) suggests that the presence of a vowel between two consonants helps explain certain cooccurrence facts—but I have not pursued this question here.

Bearing these definitions in mind, we can now examine the sorts of roots which occur in Tigrinya. The corpus used here consists of a total of 2744 roots from Bassano (1918), with the following root-template correspondences.

(27)

n	*root type*	*templatic realizations*
477	Biliterals	three-slot $C_1C_2C_2$ four-slot $C_1C_2C_2C_2$ or $C_1C_2C_1C_2$
1804	Triliterals	three-slot $C_1C_2C_3$ four-slot $C_1C_2C_2C_3$ or $C_1C_2C_3C_3$ five-slot $C_1C_2C_3C_2C_3$
463	Quadriliterals	four-slot $C_1C_2C_3C_4$ five-slot $C_1C_2C_3C_4C_4$

The one true quinquiliteral in the dictionary (with neither spreading nor regular reduplication) has been omitted from the study; see (16).

Further notes on the data are in order. Recall the generalization that identical consonants occur only at the right edge of a stem: that is, we find *sʌdʌd-* but not **sʌsʌd-*. There are 52 roots in the corpus which are superficial exceptions to this generalization, e.g. *lʌlʌy-* 'separate'. Buckley (1990) argues that these verbs necessarily involve reduplication of a shorter root plus deletion of a consonant after association to the template, e.g. *lʌylʌy* → *lʌlʌy*. These roots are treated according to this analysis, i.e. not as /lly/ but as /ly/. See also Berhane (1991: 166f) for a list of these verbs, and Greenberg (1950: 167) for a wider context.

Due to the presence of similar verbs such as *zʌlʌl-* and *zʌllʌl-* in (23), the same string of consonants may appear as a root more than once in the corpus. The duplicates have not been removed because it is considered significant that such roots occur more than once. For example, the fact that several verbs have the root /zl/ is an indication that these two consonants combine freely, whereas a single root with this combination could more easily be seen as an anomaly. Similarly, Bassano sometimes gives separate entries for a passive or causative form of a verb that may also appear as a simple entry with similar semantics.

(28) a. ˀa-rʌddʌˀ- 'persuade' *(causative in form)*
b. rʌdˀ- 'help'

(29) a. ˀam-maṡˀ- 'arrange marriage' *(caus-recip.)*
b. mʌṡˀ- 'come'

(30) a. tʌ-kʌllʌl- ‘have cataracts’ *(passive in form)*
b. kʌllʌl- ‘protect; be (partially) blinded’

Rather than second-guess his motivations, I have followed Bassano and treated these verbs as separate entries in the corpus. In other words, I have treated them in the same fashion as verbs which exist only in the causative or passive form, such as those below.

(31) ˀa-sgʌl- ‘be engrossed in work’ *(caus.)*
*sʌgʌl-

(32) ˀal-lagʌṡ- ‘ridicule’ *(causative-reciprocal)*
*lʌgʌṡ-

(33) tʌ-lʌħam- ‘spread gradually’ *(passive in form)*
*lʌħam-

Whether or not the semantic idiosyncrasies of the derived verb forms motivate this double inclusion of the roots raises issues of homonymy and polysemy well beyond the scope of this paper. At any rate it is unlikely that these roots exist in numbers significant enough to skew the results, and certainly they do not introduce spurious cooccurrences.

There are certain consonants occurring in the corpus which are excluded from the analysis. These are of two types. First, the rare segments /p, ṗ/ are found in only one verb each.

(34) a. rʌppʌs- ‘iron (clothes)’
French *repasser*
b. ṗʌṗṗʌs- ‘consecrate as bishop’ [root /ṗs/]
Greek *pappas* ‘father (bishop’s title)’

Due to their extreme rarity, the cooccurrences of these stops are simply ignored; the cooccurrence of /r/ and /s/ in /rps/, however, is included in the calculations. Second, the palatal consonants /š, j, ċ/, which as a group are uncommon in the language, particularly in verbs, have not been included in any of the statistical analysis. For example, I have not investigated whether they pattern with the coronal obstruents.

In addition, there is an assimilatory process whereby the sequence /nb/ becomes [mb]. In a typical quadriliteral verb, where the syllable pattern is *CvCCvC*, the second and third consonants occur adjacent to each other; thus when those two consonants are [mb] on the surface it is not obvious whether the nasal is underlyingly labial. There are 23 verbs of this type. Of course, if the nasal is underlyingly /m/, we have a number of new examples of the disfavored sequence of consonants from the same class. I therefore suspect that most or all of these nasals are actually /n/. There is, fortunately, a way of testing this claim: certain inflections include an /a/ between C_2 and C_3, and in the one relevant verb for which I have the necessary data, we see that the nasal is labial only when it is adjacent to the /b/.

(35)	a.	š̓ʌmbʌr-	‘mix one thing with another’
	b.	tʌ-š̓ʌnabʌr-	‘be mixed together’

Although I do not have the data required to prove it for every root, this pattern is likely to hold for the remaining 22 verbs, and I have taken the liberty of coding these nasals as /n/. Interestingly, none of these roots begins with a coronal sonorant, so that assuming /n/ as the second consonant never introduces a disfavored sequence of consonants from the same class. The third consonant, of course, is /b/; four of the roots have /l/ or /r/ as the final consonant, which as shown in (52) frequently combine with /n/. At any rate, given the small number of verbs, this assumption can have no important effect.

A similar situation holds for several roots which take a derivational prefix. These verbs can be analyzed as five-slot templates where the nasal is part of the root, or as four-slot templates with an independently attested /n/ prefix which assimilates to the root-initial /b/.

(36)	a.	tʌ-m-bʌrkʌk-	‘kneel’
	b.	tʌ-m-bʌrt̓ʌt̓-	‘be proud, strut haughtily’
	c.	ˀa-m-bahk̓ʌw-	‘yawn’
	d.	ˀa-m-bʌdbʌd-	‘fan (fire)’

I choose the latter solution, and simply omit the nasal from the root. For (36a) in particular, the word *bɨrki* ‘knee’ supports treating the root for ‘kneel’ as /brk/ rather than /nbrk/. Leslau (1941: 108) explicitly gives this analysis for (36a,b).

Leslau (1941: 106ff) identifies three consonants which serve to extend the prefixes *tʌ-* (passive) and *ˀa-* (causative): not only /n/ as illustrated in assimilated form in (36), but also /s/ and /š/. Due to templatic differences, it is easy to distinguish *prefix* + /n, s, š/ + *three-slot template* on the one hand, from *prefix* + *four-slot template* (where the first root consonant happens to be /n, s, š/) on the other. However, when the choice is between *prefix* + /n, s, š/ + *four-slot template* and *prefix* + *five-slot template* (with first root consonant /n, s, š/) there is no distinction in the overall shape of the stem (cf. Berhane 1991: 359). Along the same lines as in (36), I have uniformly treated the /n, s, š/ in such ambiguous cases as a prefixal element, and not as part of the root.

(37)	a.	tʌ-n-k̓ʌt̓k̓ʌt̓-	*root* = /k̓t̓/	‘shiver’
	b.	ˀa-s-dʌmmʌm-	*root* = /dm/	‘admire’
	c.	ˀa-š-kʌrkʌr-	*root* = /kr/	‘make turn; laugh’

This conservative decision results in fewer overall consonant cooccurrences, including those within the same class, e.g. the coronal obstruents /s, d/ in (37b), as well as those from different classes, e.g. /n, k̓/ in (37a). The number of verbs of this type is small: 44 with /n/, 13 with /s/, and just one with /š/.

3.2. Method

The basic approach taken here, as in Pierrehumbert (1993), is to compare the expected cooccurrences of each pair of consonants with the occurrences which are actually attested. The method here differs since three types of roots are included, rather than just triliterals.

The first step in this investigation was to determine the cooccurrence facts as observed in the corpus. Specifically, for each pair of consonants **a** and **b**, the cooccurrences were tallied independently for each type of root (bi-, tri-, and quadriliteral). That is, the positions in which **a** and **b** occur (call them **x** and **y**) are defined with respect to the total number of consonants in the root, as well as position within the root (I, II, III, IV). There are 10 possible pairings of root-dependent positions in a corpus that contains roots of 2, 3, and 4 consonants.

(38)

biliteral		*quadriliteral*	
I-II	ab	I-II	ab••
		II-III	•ab•
triliteral		III-IV	••ab
I-II	ab•	I-III	a•b•
II-III	•ab	II-IV	•a•b
I-III	a•b	I-IV	a••b

The cooccurrences in various roots have been kept distinct; see §5 for discussion of whether the relative frequency of consonants varies according to the different root types.

After the number of observed cooccurrences (O) for a given pair of consonants in each possible combination of positions was calculated, values were added to derive larger values for more general categories: adjacent consonants and those separated by one or two other consonants.

(39) *adjacent consonants:*

biliteral	I-II
triliteral	I-II, II-III
quadriliteral	I-II, II-III, III-IV

separated by one other consonant:

triliteral	I-III
quadriliteral	I-III, II-IV

separated by two other consonants:

quadriliteral	I-IV

In addition, values for all cooccurrences, regardless of adjacency relations (i.e. all 10 pairings), were totaled to get an overall picture.

It should be noted that this method looks only at relative position; it treats cooccurrence of **a** and **b** in /ab•/ as equivalent to cooccurrence in /•ab/. This assumption is not obviously true—initial or final position could, for example, have some special status—and the role of absolute position deserves further investigation.

The second step was to calculate the expected values for each consonant cooccurrence. First, some definitions.

(40) `A` = attested occurrences of consonant **a** in position **x**
`B` = attested occurrences of consonant **b** in position **y**
`N` = total number of roots instantiating these two positions

For each combination of two positions within a root, the expected number of pairings of **a** and **b** was calculated as follows.

(41) `A/N` = maximum likelihood estimate of the probability that **a** will occur in position **x**
`B/N` = maximum likelihood estimate of the probability that **b** will occur in position **y**
`(A/N)•(B/N)`
= probability that **a** will occur in position **x** and **b** will occur in position **y**
`E = (A/N)•(B/N)•N = (A•B)/N`
= expected number of roots with **a** in **x** and **b** in **y** from a corpus of N roots

These expected values, like the observed values discussed above, were determined independently for each type of root.

The calculation of E in (41) assumes independence of A and B, which is the null hypothesis under investigation. To the extent that the predicted value differs significantly from the observed cooccurrences, we have reason to believe that some outside factor—such as the OCP—is intervening and must be accounted for. The third step, then, was to compare the observed and expected values. For each pairing of consonants, and within each category of adjacency, the observed number of cooccurrences (O) was divided by the expected number of cooccurrences (E). A value of 1, of course, indicates that there is no factor inhibiting the cooccurrence of the consonants, i.e. that the OCP has no effect. A value of 0 arises when O=0, i.e. when the OCP effect is absolute. Values between 0 and 1 indicate varying degrees of strength of the effect. It is this varying degree—correlated with degree of similarity and degree of adjacency—which interests us here.

3.3. Results

In conformity with claims for Semitic in general, the corpus does not contain any roots with adjacent identical consonants. However, there are some roots with nonadjacent identical consonants, as well as numerous adjacent nonidentical consonants from the same place of articulation, and many nonadjacent nonidentical consonants. Statistical analysis confirms Pierrehumbert's results for Arabic regarding the role of identity and proximity: the closer and more similar the consonants, the less likely they will be found to cooccur in a root. In this section I pool data for the bi-, tri-, and quadriliterals, distinguishing the present study from Pierrehumbert's (1993), which includes only triliterals. Possible differences among these three root types are considered in §5.

3.3.1. Adjacency

As mentioned above, and observed by Greenberg (1950), same-place consonants occur with significantly less than expected frequency, but they are not absent, as identical adjacent consonants are. Some examples are given in (42) as pure roots, rather than as stems.

(42) *Some roots with adjacent homorganic consonants*

sd	'send'
zt̓	'lean'
lbm	'be rational, shrewd'
lk^{w}k	'seal with a cover'
gss̓	'reproach, warn, correct'
wsd	'take (along)'
bt̓s	'break rope, wire'
szy	'not come true'
ss̓y	'waste away from disease'
ħrnk	'snore'
mnt̓s	'tear out by roots'

The table in (43) shows the observed occurrences of such adjacent homorganic consonants in the corpus (O); the expected occurrences based on the frequency of relevant consonants in the necessary positions (E); and the degree to which cooccurrence is actually permitted (O/E).[3]

[3] In the tables here, figures for E have been rounded to whole numbers, but the exact values were used in calculating O/E.

(43) *Cooccurrences of adjacent homorganic consonants*

	O	*E*	*O/E*
guttural	0	78	0.00
velar	1	125	0.01
coronal obstruent	65	242	0.27
coronal sonorant	27	261	0.10
labial	2	132	0.02

Thus only one root contains two velars in adjacent positions—specifically, /lk^wk/ in (42)—but the overall frequency of velar consonants would lead us to expect 125 such roots. It can be observed that the effect is weaker for coronals than for the other classes; see also below.

Homorganic consonants are more common when nonadjacent than when adjacent, which the standard approach in (22) fails to capture in a principled way. Roots of this type are given in (44), and the numerical summary in (45). By definition, no biliterals appear in this group.

(44) *Some roots with nonadjacent homorganic consonants*

t̓mt	‘stare at’
fsm	‘become pale or discolored’
drt	‘delimit a field’
bg^wm	‘be sly, taciturn’
ħsʔ	‘be dry (esp. of hair)’
zmd	‘become related by marriage’
t̓ys	‘repopulate a region with former residents’
mlbs	‘be weak and incapable of work’
wslt	‘lie, cheat; miss a target’

The term ‘nonadjacent’ here includes only consonants separated by exactly one consonant. Chi-square tests show excellent significance for the adjacent and separated-by-one cases ($p \leq 0.001$), but no significance for the separated-by-two cases ($p \leq 0.250$): the number of tokens for the last category is too small. Data from this much smaller group of pairings separated by two consonants (i.e. I-IV in quadriliterals) are excluded in (45) to (49), but are consistent with these results.

(45) *Cooccurrences of nonadjacent homorganic consonants (separated by one consonant)*

	O	*E*	*O/E*
guttural	6	49	0.12
velar	10	73	0.14
coronal obstruent	112	162	0.69
coronal sonorant	106	140	0.76
labial	18	68	0.26

Since identical consonants are also homorganic, they are included in (45). The table in (46) compares more directly the strength of the OCP effect for these two classes of cases, divided by place of articulation. The lower the value of O/E, the stronger the inhibiting effect.

(46) *O/E values for homorganic consonants by adjacency*

	adjacent	*nonadjacent*
guttural	0.00	0.12
velar	0.01	0.14
coronal obstruent	0.27	0.69
coronal sonorant	0.10	0.76
labial	0.02	0.26

For each place of articulation, the effect is weaker when nonadjacent; that is, the value O/E is higher for the nonadjacent pairings. It is also the case that the OCP effect is weaker for coronals in both situations. It should be noted in addition that since biliterals are inherently excluded from the nonadjacent class, there is a possible confounding of the effect of adjacency with some difference between biliterals and other roots. However, the data in §5 indicate that biliterals and triliterals pattern in the same way.

3.3.2. Identity

Adjacent identical consonants are absolutely prohibited (so this again excludes biliterals), but the following roots illustrate the fact that identical consonants do occur in nonadjacent positions.[4]

[4] Some of these roots have known historical origins in roots without identical

(47) *Some roots with nonadjacent identical consonants*

sls	'plow a field the third time'
lʔl	'raise, lift off the ground'
trt	'tell stories, old traditions'
dndw	'threaten to hit'
mslm	'convert to Islam'

Only 12 such roots exist in the corpus, but just 125 are expected in the first place, since the requirement of full identity is quite stringent.

(48) *Nonadjacent cooccurrences of identical consonants*
(separated by one consonant)

	O	*E*	*O/E*
guttural	0	16	0.00
velar	1	14	0.07
coronal obstruent	7	30	0.23
coronal sonorant	2	43	0.05
labial	2	22	0.09

The table in (48) excludes six roots with identical labiovelars in the pattern $C_iC_jC_iC_k$, which appear to be modified reduplications; see below for discussion.[5] Within the nonadjacent set, **nonidentical** consonants of the same articulatory class are far more common than those which are identical (Pierrehumbert 1993).[6] This is not predicted at all by the place OCP.

consonants. For example, /sls/ is the root for 'three'; earlier this was /śls/ (as in Ge'ez: Leslau 1987), but in Tigrinya the two fricatives merged. (However, other branches of Semitic also have identical obstruents, e.g. Arabic /θlθ/, Hebrew and Akkadian /šlš/.) Similarly, the root /mslm/ is from the Arabic noun *muslim*, where the /m/ belongs to a prefix and the original root is /slm/, with just one labial.

[4] Without this omission, the figures in (48) for velars are O=7, E=14, and quite aberrant O/E=0.51.

[6] The 'nonadjacent homorganic' figures in (45) are the sum of 'nonadjacent identical' in (48) and 'nonadjacent nonidentical' (but, of course, homorganic) in (49).

(49) *Nonadjacent cooccurrences of nonidentical consonants (separated by one consonant)*

	O	*E*	*O/E*
guttural	6	33	0.18
velar	3	59	0.05
coronal obstruent	105	132	0.80
coronal sonorant	104	97	1.07
labial	16	46	0.35

As usual, among the coronals the prohibition is weaker. In fact, no prohibitory effect is found for nonadjacent coronal sonorants, where O/E>1.

Important subregularities exist within the major articulatory classes (for Arabic, cf. Yip 1989, Padgett 1992, Pierrehumbert 1993, McCarthy 1994). The guttural class, for example, consists of the laryngeals /h, ʔ/ and the pharyngeals /ħ, ʕ/. While a laryngeal occasionally combines with a pharyngeal, there are no cooccurrences at all within either of these subclasses. This gap is particularly striking in the case of the pharyngeals, which are rather common in the corpus.

(50) *Cooccurrences of gutturals (regardless of adjacency)*

	O	*E*	*O/E*
laryngeal + pharyngeal	6	49	0.14
two laryngeals	0	8	0.00
two pharyngeals	0	78	0.00

This greater inhibition is due to relative similarity: the more similar the consonants are, the less they cooccur. So while /h, ʔ, ħ, ʕ/ share the property of being gutturals and their cooccurrence is consequently disfavored, the subset /ħ, ʕ/ shares the further property of being pharyngeals, and the prohibition on cooccurrence is absolute in this corpus.

Among coronals, the very fact that the sonorants and obstruents are placed in separate classes is due to the same notion of similarity. Perhaps due to the large size of the coronal class and the many distinctions that exist among its members, the central [±sonorant] dichotomy is strong enough to eliminate any OCP effect (cf. Greenberg 1950: 162f, Padgett 1992). The following table illustrates the lack of an effect across the two classes in Tigrinya.

(51) *Cooccurrences of a coronal sonorant and obstruent*

	O	*E*	*O/E*
adjacent	744	625	1.10
nonadjacent	339	290	1.17

Compare these large values for O/E to the lower values in (46). Within each of these coronal classes, we find further effects of relative identity. Among the coronal sonorants, /n/ combines rather freely with /l, r/, but the liquids never cooccur with each other (cf. Greenberg 1950: 172f, Pierrehumbert 1993).

(52) *Cooccurrences of coronal sonorants (regardless of adjacency)*

	O	*E*	*O/E*
one each of /n, l/	46	81	0.57
one each of /n, r/	85	99	0.86
one each of /r, l/	0	91	0.00

This result suggests that the most salient feature among the sonorants is [±nasal], splitting the members into the two classes /n/ and /l, r/. Within either class the cooccurrence restriction is absolute in effect,[7] but across the classes the effect is weak.

Among the coronal obstruents, the important dichotomy is defined by [±continuant]. The fricatives /s, z, ṡ/ combine rather freely with the stops /t, d, ṫ/, but within these classes cooccurrence is much less frequent (cf. Padgett 1992, Pierrehumbert 1993).

[7] Strictly speaking, the effect is not absolute within the /n/ class, because there are two quadriliterals which contain two instances of /n/: *tʌ-mʌnk̇ʌn-* 'be cunning' and *tʌ-kʌnawʌn-* 'succeed'. These positions, II and IV, are the typical locations of those identical consonants that do occur in quadriliterals, and seem to involve a kind of semi-reduplication. There are also two roots with identical liquids: the common root /lʕl/ 'on, above' (originally from the preposition *la-* plus the root /ʕly/ 'up, high': Leslau 1987: 304) and the unusual quinquiliteral *gʌrʌngʌr-* (16).

(53) *Cooccurrences of coronal obstruents (regardless of adjacency)*

	O	E	O/E
fricative and stop	178	216	0.82
two fricatives	9	103	0.09
two stops	9	112	0.08

See (61) below for further data on coronal cooccurrence effects.

Finally, within the velar group, there is a striking difference between the plain velars /k, g, k̇/ and the labiovelars /k^{w}, g^{w}, k̇^{w}/. Cooccurrences among the labiovelars are far more common than among the plain velars.

(54) *Some roots with multiple velars*

lk^{w}k	'seal with a cover'
g^{w}nk^{w}	'transport (grain) to threshing floor'
g^{w}ħnk̇^{w}	'have chin joined to chest'
k̇^{w}yk̇^{w}	'quarrel'

The only root which contains two plain velars is /grngr/ in (16), which is excluded from the analysis. Even if this root were added to the calculation for the table in (55), it would not change the stark difference between the relatively frequent cooccurrences of labiovelars (O/E=0.48), as opposed to cases where at least one plain velar is involved (less than 0.02).

(55) *Cooccurrences of velars (regardless of adjacency)*

	O	E	O/E
plain velar + labiovelar	2	88	0.02
two plain velars	0	96	0.00
two labiovelars	10	21	0.48

While a rather large number of labiovelar pairs exist, they are all in nonadjacent positions, where there seems to be no OCP effect at all.

(56) *Cooccurrences of labiovelars*

	O	*E*	*O/E*
adjacent	0	11	0.00
nonadjacent	10	10	1.05

Excluded from (48) above, but included here in (56), are six suspicious cases of cooccurring labiovelars. They all involve positions I and III in quadriliterals, suggesting a type of semi-reduplication. Five of them have /r/ in position II: /kʷrkʷm, kʷrkʷħ, kʷrkʷˀ, gʷrgʷħ, gʷrgʷc̓/, while the sixth has /r/ in position IV: /k̓ʷnk̓ʷr/. It is likely that most or all of these derive historically from reduplicated biliterals by substitution of /r/ (or /n/) in position II, or from a triliteral with infixation of the first consonant between the last two. Both possible origins are attested for Arabic, e.g. *qafqaf/qarqaf, ṭarib/ṭarṭab* (Fleisch 1968: 128f). Another scenario can be suggested for /gʷrgʷħ/ 'empty a container', which resembles /drgʷħ/ 'pour liquid from a vessel': here the initial /d/ has apparently been replaced to create a semi-reduplication. A plausible connection also exists between /k̓ʷnk̓ʷr/ 'have sunken eyes' and /nk̓ʷr/ 'be blinded'. Whatever the exact origin of each root, the original forms were probably consistent with the OCP.

If these six roots are omitted from (56), overall O/E for labiovelars is reduced to 4/21 = 0.19, mitigating their difference from plain velars, but certainly not eliminating it. The remaining asymmetry—as well as the existence of the six aberrant roots—is perhaps due to the fact that the labiovelars are not inherited from Proto-Semitic, and seem attributable from the Cushitic substrate in Ethiopia (cf. Dillman and Bezold 1907: 50f). If the source language(s) for borrowed words did not enforce a place OCP on roots, then the resulting roots will bring with them violations of the native Semitic pattern. As we have seen, this prohibition is not absolute, and a certain number of such violations can be tolerated.

An additional point of interest regarding the labiovelars is their status relative to the labials. In Arabic, it is well known that consonants with secondary pharyngealization do not resist combination with the pharyngeal class (Greenberg 1950, McCarthy 1994, Pierrehumbert 1993). A similar nonequivalence of secondary and primary articulations exists for labiality in Tigrinya, since labiovelars occur quite often with labials.

(57) *Cooccurrences of a labial and labiovelar*

	O	*E*	*O/E*
adjacent	71	85	0.84
nonadjacent	98	60	1.64

There is apparently a weak effect in adjacent position, but certainly no effect for nonadjacent tokens. It is not clear whether this difference is related to the considerable adjacency asymmetry for labiovelars as a class, seen in (56). Further research on the history of labiovelars in Tigrinya may elucidate these patterns.

4. The Similarity Model

As Pierrehumbert (1993) notes, the standard approach to Semitic consonant cooccurrence restrictions requires two constraints: the total OCP, acting on adjacent segments, is absolute, while the place OCP, acting on potentially nonadjacent articulator nodes, is not absolute. The account is therefore not unified. In addition, the gradient effect of the place OCP, based on relative adjacency and relative similarity, is unexplained. Pierrehumbert's data for Arabic, as well as the data for Tigrinya presented in §3, show that the extent to which the cooccurrence of two consonants is disfavored depends on how close they are (degree of adjacency) and how many features they share (degree of identity); see also Greenberg (1950). It appears, then, that the OCP is not an iron-clad principle, but is instead a matter of degree.[8]

Pierrehumbert (1993) proposes that there is only a place OCP which targets (tier-adjacent) identical articulator nodes, but whose strength of effect is proportional to the similarity of the consonants which have those identical nodes. Similarity increases with proximity and featural identity, and these effects are cumulative. The case of maximal similarity—where the consonants are both adjacent and fully identical in their features—is prohibited absolutely. The extreme case of enforcement derives the effect of the total OCP, but is treated formally as a subcase of the place OCP. Thus by incorporating gradience into the analysis, the disunity of the standard approach (both place and total OCP) can be eliminated.

[8] See Berkley (1994a,b) for similar arguments based on English consonant cooccurrences.

This formulation relates to a more general notion of perceptual similarity, which is independently known to depend on the presence of intervening material (cf. Zechmeister and Nyberg 1982) and to be analyzable in terms of feature sets (cf. Tversky 1977). Analogies for Semitic root consonant strings are given in (58), using diagrams broadly inspired by those given by Tversky.

(58) a. *Adjacent, identical* (prohibited by OCP); cf. `bbk`

b. *Adjacent, nonidentical* (disfavored); cf. `bfk`

c. *Nonadjacent, identical* (disfavored); cf. `bkb`

d. *Nonadjacent, nonidentical* (mildly disfavored); cf. `bkf`

The smiling face in (58a) is more easily recognized as similar to an identical face when that second face is adjacent to it, in essence facilitating comparison. In the same way, a /b/ is easily recognized as identical to another /b/ when it is adjacent to it—and so this root type is ruled out. When similarity is actually lessened by a change in features—whether a change in expression for the face in (58b), or a change in continuancy and voicing for /f/—the two objects are less offensive to any principle against adjacent identical items. But a reduction in **perceived** similarity is achieved by moving the objects further apart, with an intervening object, whether another kind of "face" (58c) or another consonant. The two effects combine in a case like (58d), where nonidentical, nonadjacent objects are the least likely to be perceived as similar.

Pierrehumbert (1993) proposes a simple mathematical formula for calculating the degree of Similarity according to featural identity.[9] It

---

[9] For a revised approach, not pursued here, see Frisch, Broe, and Pierrehumbert (1995). See also Pierrehumbert (1994) for more general discussion.

depends on the distinctive features that are assumed for the consonants in question—more specifically, on how many of these features of the consonants are the same, and how many are different.

(59) a. Same = number of shared features
b. Different = number of features which differ
c. Similarity (S) = Same/(Same + Different)

The advantage of this formulation is that it produces a result that ranges from 0, if no features are the same: 0/(0+*n*); to 1, if no features are different: *n*/(*n*+0). I propose an additional metric of Distinctness, which is simply the difference between Similarity and 1.

(60) Distinctness (D) = 1 – S

Ideally, this figure will stand in rough correspondence to O/E. That is, as Similarity increases, Distinctness (D) decreases, and the greater strength of the OCP effect should also lead to a lowering of the value for O/E.

This approach provides a means of testing proposals for feature specifications (cf. Pierrehumbert 1993). Here I pursue an interesting confirmation of recent arguments that laryngeal features are privative. Consider the following fricative-stop pairings according to laryngeal articulation.

(61) *Cooccurrences of coronal obstruents*
(regardless of adjacency)

	O	*E*	*O/E*
voiceless /s, t/	39	70	0.56
voiced /z, d/	8	43	0.19
ejective /s̉, t̉/	4	33	0.12

The voiceless consonants cooccur much more freely than either the voiced or ejective pairs. Yet at first glance it seems that each of these pairs is comparable, differing only in [continuant]. To capture the special status of the voiceless pair, /s, t/ must count as less similar than either /z, d/ or /s̉, t̉/.

There is considerable evidence from a wide range of languages and phonological phenomena that the features which define these pairs, i.e. [voiced] and [constricted glottis], are privative (Mester and Itô 1989, Lombardi 1991, 1995); that is, they exist only in their positive values, and there is no formal object corresponding to [–voiced] or [–constricted glottis]. A plain voiceless obstruent, then, bears no laryngeal features at all.

(62)

	s	z	ṡ	t	d	ṫ
[Coronal]	+	+	+	+	+	+
[continuant]	+	+	+	–	–	–
[voiced]	0	+	0	0	+	0
[constricted glottis]	0	0	+	0	0	+

In order to capture this difference in feature marking, the determination of similarity must ignore the joint lack of a feature (cf. Yip 1989).

(63)

Same	[+ +], [– –]
Different	[+ –], [+ 0], [– 0]
Neither	[0 0]

Under these assumptions, pairs of consonants in which both possess the same privative feature ([voiced], [constricted glottis]) receive a greater relative value for Same and therefore greater S. Where both members of a pair lack a feature, that shared property is not counted and a smaller value for S results.

(64)

	Same	*Different*
voiceless /s, t/	[Cor]	[cont]
voiced /z, d/	[Cor]	[cont, voiced]
ejective /ṡ, ṫ/	[Cor]	[cont, cg]

Using the formulas given in (59) and (60), we arrive at the following values for Similarity and Distinctness. The values for O/E in (61) are repeated below as well.

(65)

	S	*D*	*O/E*
voiceless /s, t/	0.50	0.50	0.56
voiced /z, d/	0.67	0.33	0.19
ejective /ṡ, ṫ/	0.67	0.33	0.12

While the correspondence of D and O/E is far from exact, the overall pattern is the same: the voiceless pair is higher than the other two, which are (roughly) the same as each other. The important point is that the achievement of this asymmetry depends on accepting the privative status of laryngeal features, and provides support for that conclusion from an unexpected source.

Although they cannot be pursued here, interesting questions arise from the nature of the calculations required to correctly predict the strength of the OCP effect based on the featural representations (see Frisch, Broe, and Pierrehumbert 1995). In particular, certain features seem to carry more weight than others. For example, the table in (52) suggests that a difference in [nasal] entails greater distinctness than a difference in [lateral], thereby permitting freer cooccurrence of /n/ with /l/ or /r/ than of the two liquids together (cf. Pierrehumbert 1993). The special status of [lateral] in the system, as a feature which serves to distinguish only this pair of sounds, may be the ultimate explanation; or it might be captured in approaches to the representation of sonorants, such as the Spontaneous Voice node of Rice and Avery (1991). I leave resolution of such questions for future research.

5. The Role of Root Length

In this section I consider an additional hypothesis which is testable due to the inclusion of more than just triliterals in the corpus. Specifically, is there evidence that perceived similarity of adjacent consonants decreases as the total number of consonants in the root increases? We might call this the ‘distraction’ factor: just as an intervening item reduces perceived similarity, perhaps the presence of flanking items serves a similar function.[10]

This hypothesis can be meaningfully tested in Tigrinya only for the coronals, because the other classes cooccur too infrequently in all root types: see (43) and (45). Among the coronal sonorants, there is a significant difference between tri- and quadriliterals, but not between bi- and triliterals. Since biliterals are included in the comparison, only adjacent pairs are considered for the longer roots.

[10] Root length is clearly relevant to significance. Given 27 consonants, there are $27^2 = 729$ possible biliterals, $27^3 = 19{,}683$ possible triliterals, and 27^4 = over 14 million possible quadriliterals. Thus the absence of a particular cooccurrence among the biliterals is far more significant than the absence of a cooccurrence among the triliterals.

(66) *Cooccurrences of adjacent coronal sonorants*

	O	*E*	*O/E*
biliteral	0	15	0.00
e.g. *nr* (absent)			
triliteral	8	178	0.04
e.g. *knr, nrk* (very rare)			
quadriliteral	19	67	0.28
e.g. *bknr, bnrk, nrbk* (uncommon)			

Notice that both bi- and triliterals show a strong OCP effect, while quadriliterals show a much weaker effect. A possible explanation is that the presence of the other consonants creates a distraction effect which reduces perceived similarity, and thus reduces the strength of the OCP. However, it is surprising that the same effect does not distinguish the bi- and triliterals.

Among the coronal obstruents, consideration of the full data does not yield the same pattern. Rather, we find very similar O/E values for each root type.

(67) *Cooccurrences of adjacent coronal obstruents*
('quadriliteral' = all pairings: I-II, II-III, III-IV)

	O	*E*	*O/E*
biliteral	9	30	0.30
triliteral	44	168	0.26
quadriliteral	12	43	0.28

These data suggest that there is no distraction effect. However, the pattern for the obstruents closely parallels that for the sonorants if we omit the pairing of the last two consonants (III-IV) in the quadriliterals. This pairing is strikingly aberrant, as demonstrated by its divergent O/E value.

(68) *Cooccurrences of adjacent coronal obstruents*
in quadriliterals

	O	*E*	*O/E*
pairing I-II	4	8	0.50
pairing II-III	4	8	0.49
pairing III-IV	4	27	0.15

In other words, the pattern *bkts* is proportionally far rarer than *tsbk* or *btsk*. The explanation for this asymmetry is unknown, but possibly diachronic. It is interesting to note that the absolute numbers of observed cooccurrences are the same; but in III-IV the expected numbers are much greater, because for quadriliterals in general, position IV contains a large number of coronal obstruents. This in turn is perhaps because some coronal-obstruent suffix has been incorporated into many historically triliteral roots, but subject to the OCP, so that its current distribution is skewed in favor of roots which do not contain another coronal obstruent, at least in adjacent position. At any rate, if we exclude the pairing III-IV from the quadriliteral data, we arrive at a pattern more similar to that for the sonorants in (66).

(69) *Cooccurrences of adjacent coronal obstruents*
('quadriliteral' = pairings I-II, II-III, but not III-IV)

	O	*E*	*O/E*
biliteral	9	30	0.30
triliteral	44	168	0.26
quadriliteral (nonfinal)	8	16	0.49

As with the sonorants, the bi- and triliterals are roughly the same, while the quadriliterals show much greater freedom of cooccurrence.

While the modification of the data in (69) produces a similar pattern to that in (66), which together might be taken to suggest a distraction effect in the quadriliterals, the fact remains that there is no such difference between the bi- and triliterals. Since the quadriliterals introduce considerable complications—they are in some cases reanalyses of a triliteral plus an affix, or involve semi-reduplications—there are many other potential explanations for their differences from the shorter roots. Based on the more straightforward bi- and triliteral data, there does not in fact appear to be any distraction effect.

5. Conclusion

I have shown that the patterning of root consonants in Tigrinya obeys the same restrictions on homorganic cooccurrences found in Arabic and other Semitic languages. In particular, the data support the conclusion of Pierrehumbert (1993) that the OCP which holds of place features relies on a gradient notion of similarity. The case of maximal similarity (adjacent and identical) is ruled out; identical consonants obscured by nonadjacency are permitted in small numbers; nonidentical (homorganic) consonants are permitted with some freedom if adjacent, and with considerable freedom if separated by one or more consonants. Comparison of bi- and triliteral roots also suggests that the number of consonants in a root has no bearing on the nature of consonant cooccurrences found in the root.

References

Bassano, Francesco da. 1918. *Vocabolario tigray-italiano*. Casa editrice italiana di C. de Luigi, Rome.

Bender, M.L. 1978. Consonant cooccurrence restrictions in Afroasiatic verb roots. *Atti del secondo congresso internazionale di linguistica camito-semitica.* (Quaderni di Semitistica 5.) Istituto de Linguistica e di Lingue Orientale, Università di Firenze, pp. 9–19.

Berhane, Girmay. 1991. *Issues in the phonology and morphology of Tigrinya.* Dissertation, Université du Québec à Montréal.

Berkley, Deborah Milam. 1994a. Variability in Obligatory Contour Principle effects. *Papers from the 30th Regional Meeting of the Chicago Linguistics Society. Vol. 2: Parassession on Variation in Linguistic Theory*, pp. 1–12.

Berkley, Deborah Milam. 1994b. The OCP and gradient data. *Studies in the Linguistic Sciences* 24, 59–72.

Buckley, Eugene. 1990. Edge-in association and OCP 'violations' in Tigrinya. *Proceedings of the Ninth West Coast Conference on Formal Linguistics*, pp. 75–90.

Dillman, August, and Carl Bezold. 1907. *Ethiopic grammar*. (Translated by James A. Crichton.) Williams & Norgate, London.

Fleisch, Henri. 1968. *L'arabe classique: esquisse d'une structure linguistique*. Dar el-Machreq Éditeurs, Beyrouth.

Frisch, Stefan, Michael Broe, and Janet Pierrehumbert. 1995. The role of similarity in phonotactic constraints. Ms., Northwestern University.

Goldsmith, John. 1976. *Autosegmental phonology*. Dissertation, MIT. Published 1979 by Garland, New York.

Greenberg, Joseph. 1950. The patterning of root morphemes in Semitic. *Word* 6, 162–181.

Leben, William. 1973. *Suprasegmental phonology*. Dissertation, MIT. Published 1980 by Garland, New York.

Leslau, Wolf. 1941. *Documents tigrigna (éthiopien septentrional): grammaire et textes*. Librairie C. Klincksieck, Paris.

Leslau, Wolf. 1987. *Comparative dictionary of Ge'ez (classical Ethiopic)*. Harrassowitz, Wiesbaden.

Lightner, Theodore M. 1973. Against morpheme structure rules and other things. In M. Kenstowicz and C. Kisseberth (eds.), *Issues in phonological theory*. (Janua linguarum, series maior 74.) Mouton, The Hague, pp. 53–60.

Lombardi, Linda. 1991. *Laryngeal features and laryngeal neutralization*. Dissertation, University of Massachusetts, Amherst.

Lombardi, Linda. 1995. Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory* 13, 39–74.

McCarthy, John J. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12, 373–418.

McCarthy, John J. 1986. OCP effects: gemination and anti-gemination. *Linguistic Inquiry* 17, 207–263.

McCarthy, John J. 1988. Feature geometry and dependency: a review. *Phonetica* 43, 84–108.

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In P. Keating (ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III*. Cambridge University Press, pp. 191–233.

Mester, R. Armin. 1986. *Studies in tier structure*. Dissertation, University of Massachusetts, Amherst.

Mester, R. Armin, and Junko Itô. 1989. Feature predictability and underspecification: palatal prosody in Japanese mimetics. *Language* 65, 258–293.

Padgett, Jaye. 1992. OCP subsidiary features. *Proceedings of the Annual Meeting of the North Eastern Linguistics Society* 22, 335–346.

Pierrehumbert, Janet. 1993. Dissimilarity in the Arabic verbal roots. *Proceedings of the Annual Meeting of the North Eastern Linguistics Society* 23, 367–381.

Pierrehumbert, Janet. 1994. Knowledge of variation. *Papers from the 30th Regional Meeting of the Chicago Linguistics Society. Vol. 2: Parassession on Variation in Linguistic Theory*, pp. 232–256.

Rice, Keren, and Peter Avery. 1991. On the relationship between laterality and coronality. In C. Paradis and J.-F. Prunet (eds.), *The special status of coronals* (Phonetics and Phonology 2). Academic Press, San Diego, pp. 101–124.

Sagey, Elizabeth. 1986. *The representation of featrures and relations in nonlinear phonology*. Dissertation, MIT. Published 1991 as *The representation of features in non-linear phonology: The articulator node hierarchy*, Garland, New York.

Tversky, A. 1977. Features of similarity. *Psychological Review* 84, 327–352.

Yip, Moira. 1988. Template morphology and the direction of association. *Natural Language and Linguistic Theory* 6, 551–577.

Yip, Moira. 1989. Feature geometry and cooccurrence restrictions. *Phonology* 6, 349–374.

Zechmeister, E. and S. Nyberg. 1982. *Human memory*. Brooks/Cole, Monterey.

Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305
gene@unagi.cis.upenn.edu