Syllables and Moras in Arabic

Paul Kiparsky

6.1. CV, VC-, and C-dialects

Some of the most salient differences among Arabic vernaculars have to do with syllable structure. This study focuses on the syllabification patterns of three dialect groups, (1) VC-dialects, (2) C-dialects, and (3) CV-dialects, and argues that they differ in the licensing of semisyllables, moras unaffiliated with syllables and adjoined to higher prosodic constituents. The analysis provides some evidence for a constraint-based version of Lexical Phonology, which treats word phonology and sentence phonology as distinct constraint systems that interact in serial fashion.

VC-dialects include the dialects of Syria, Lebanon, Palestine, Iraq, and Turkey (Blanc 1953, Erwin 1963, Cowell 1964, Grotzfeld 1965, Palva 1966, Jastrow 1978, Behnstedt 1994), Bedouin and Bedouin-type dialects such as Bani-Hassan of Jordan (Irshied and Kenstowicz 1984), the Hijazi dialects of Central Arabia (Jastrow 1980a), and the dialects of eastern Libya (Owens 1984: 12ff., Mitchell 1993: 85, 88), and two groups of Egyptian dialects, spoken in the easternmost part of the Delta and in Upper Egypt approximately to Asyut (Woidich 1980: 207, Behnstedt and Woidich 1985).


I would like to thank Andrew Garrett, Larry Hyman, John McCarthy, and Michael Redford for comments and discussion, and Albert Borg and Mauwel Mihal for their expert counsel on Maltese. The final version has benefited from the comments of two readers. After this chapter was written, I became aware that Högstrom (1997) (who in turn refers to Potter 1994) had proposed more or less the same idea of unsyllabified moras for Mohawk and Passamaquoddy.
which have been analyzed both as complex clusters and as sequences of syllables with consonantal nuclei. Certain Bedoin-type dialects also seem to belong in this group, as does the Maltese language.

CV dialects constitute a distinctive group, comprising the majority of the dialects of Egypt, including Cairo, most of the Delta, the oases of the Libyan desert, and Middle Egypt (Woidich 1980, Behnstedt and Woidich 1985). They correspond to Broselow’s (1992) onset dialects.1

VC- and C-dialects often coexist as distinct speech registers or sociolects within a basic regional dialect. I will argue that they have the same lexical syllabification and differ mainly in their postlexical phonology. The CV-dialects, with a distinct lexical syllabification, differ more fundamentally from both.

The table in (1) illustrates the distribution of diagnostic properties for 15 Arabic dialects and for the Maltese language. Column G is included to show that the syllabification isoglosses do not bundle with the major word stress isogloss. The Egyptian data are from Behnstedt and Woidich (1985), and those from other dialects are from Fischer and Jastrow (1980), Mitchell (1993), and the additional sources indicated in (2).2

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>kalib</td>
<td>katāḥit</td>
<td>(ʔ)ktāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>i.</td>
<td>kalb</td>
<td>katab</td>
<td>ktāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>ii.</td>
<td>kalb</td>
<td>katab</td>
<td>ktāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>iii.</td>
<td>kalib</td>
<td>katāḥit</td>
<td>(ʔ)ktāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>iv.</td>
<td>alab</td>
<td>alāb</td>
<td>alāb</td>
<td>yidāl(i)mān</td>
<td>bābha</td>
<td>maktāba</td>
<td></td>
</tr>
<tr>
<td>v.</td>
<td>kalb</td>
<td>katāḥit</td>
<td>(ʔ)ktāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>vi.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>b.</td>
<td>kalb</td>
<td>kathāb</td>
<td>kathāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>i.</td>
<td>kalb</td>
<td>kathāb</td>
<td>kathāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>ii.</td>
<td>kalb</td>
<td>kathāb</td>
<td>kathāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>iii.</td>
<td>kalb</td>
<td>kathāb</td>
<td>kathāb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>c.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>i.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>ii.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>iii.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>iv.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>v.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>vi.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
<tr>
<td>vii.</td>
<td>kalb</td>
<td>kalb</td>
<td>kalb</td>
<td>yikītbu</td>
<td>ikālmun</td>
<td>bābha</td>
<td>maktāba</td>
</tr>
</tbody>
</table>

A summary of the crc
Phrase-final -CC el (columns A-B). They conditions that vary accord
kataa' (CV-dialects),
Rīsma rīsma, rīsma 'na
(kalib, kataa') or perr
Phrase-initial onset
accordingly allow the
syllables to reduce or
are often broken
preceded by a glottal
of Onset); for exam
(ʔ)iyyaar 'donkey' (sii
Geminates are incl
fact that will be of spec
assimilation, initial ger
resolved by epenthesis
llandān/ llandān, ʔilla
as sequences of syllable-structure also seem to play a role in the majority of the cases of the Libyan and Wadiel dialects. They might be, in some instances, part of a greater dialectal structure or sub-dialects.

The CV-dialects have the same lexical and phonological properties as the VC-dialects for the major word stress (1985), and those described for Arabic (Mitchell 1993), and

<table>
<thead>
<tr>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>her door</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
<tr>
<td>bâbba</td>
<td>maktaba</td>
</tr>
</tbody>
</table>

A summary of the cross-dialectal generalizations follows.

**Phrase-final -CC clusters** occur unrestrictedly only in CV- and C-dialects (columns A–B). They can be broken up by an epenthetic vowel, under conditions that vary according to style and dialect: for example, /kata'b/ katabit, kata'b (CV-dialects), kâbî (C-dialects) 'I wrote', /qāl/ qâl, qâl 'food', /fâm/ fâm, fâm 'name'. The VC-dialects either permit no -CC clusters (ka'b, katabit) or permit them only with falling sonority (ka'b, katabit).

**Phrase-initial onset CC- clusters** are allowed in VC- and C-dialects, which accordingly allow the pan-Arabic process that deletes high vowels in open syllables to reduce even initial CC- to CC- (column C). The resulting clusters are often broken up by a prothetic vowel (which in turn is phonetically preceded by a glottal stop, in satisfaction of Arabic's undominated ranking of Onset); for example, (/i/)sâlah, 'plowshare', (/i/)kla'ab, (/i/)kla'ab 'dogs', (/i/)h'mâr 'donkey' (sâlah, kila'ab, h'mâr in CV-dialects).

**Geminates** are included in the class of permissible initial CC- clusters, a fact that will be of special significance in my analysis. Most often arising from assimilation, initial geminates occur only in VC- and C-dialects. They can be resolved by epenthesis in the same way as other onset clusters; for example, /lândan/ lândan, lândan 'to London', /lã'ã/ lã'ã 'the tea'.

(2) Key to the dialects shown in (1)

a. VC-dialects
   i. Šarqiyâ dialects (easternmost part of Delta)
   ii. Upper Egypt south of Asyût
   iii. Iraq (Baghdad: Erwin 1963)
   iv. Syria (Soukhnâ: Behnstedt 1994)
   v. Syria (Damascus: Cowell 1964, Grotfeldt 1965)
   vi. Eastern Libya (Owens 1984)

b. C-dialects
   i. Tunisia (Marçais 1977, Singer 1980)
   ii. Rafah, southern Hijaz (Prochazka 1988: 32, 153, 162, 179, 185, 198, 200)
   iii. Maltese (Aquilina 1959)

c. CV-dialects
   i. Cairo
   ii. Rosetta
   iii. Damiata (Dumyât)
   iv. ilBalqâriâ oasis
   v. alFarâfîra oasis
   vi. adDâxîla oasis
   vii. Middle Egypt, Upper Egypt to Asyût
Medial -CCC- clusters are broken up as -GCC- in VC-dialects and as -CGC- in CV-dialects; for example, Iraqi ᵍʰliːla, Cairene ḫullu(h), Moroccan qṣṭṭu 'you (m.) said to him'.

"Metathesis" of medial -CIC- to -CIC- occurs only in VC-dialects (column D); for example, ʰ-kiṭīb-u yikību 'they write'. CV-dialects always retain -CGC- (yikību), and C-dialects simply drop the vowel in the corresponding cases (yikību). Woidich (1980: 212) and independently Broselow (1992: 35) noticed that dialects have metathesis only if they have epenthesis of the -CCC- → -CGC- type, and the correlation with the other VC-dialects traits can be seen for Egyptian in Behnstedt and Woidich 1985: maps 59, 66, 67. As Woidich and Broselow point out, this is evidence for interpreting metathesis as a composite process consisting of medial syncope followed by insertion of an epenthetic vowel into the resulting -CGC- cluster (-CGC- → -CCC- → -CIC-).

Desonorization of word-final -VCR-, -VVR occurs only in CV-dialects. For example, in North Yemen the CV-NV- isogloss coincides with the incidence of glottalization and devoicing (-VCR-, -YR) (Behnstedt 1985: 14, 48, 58). In Egypt, final glottalization occurs in a subset of CV-dialects; for example, [kɑʔ]ḥ, moʔz] (Behnstedt and Woidich 1985, maps 41–43).

High vowel deletion occurs after geminates only in the VC- and C-dialects (column E); for example, ʰ-kiṭīb-u (y)i-kal(l)mu 'they talk to someone', ʰ-sakkīb-u (y)sak(ī)r, 'they shut', ʰ-sallīm-u (y)i-Tal(l)mu 'they teach'. All CV-dialects retain the vowel (y)kal(l)nu, y(y)tal(l)nu, y(s)akkiru (or yikallimu etc.).) In most dialects, the geminate is then shortened, but retention of a quantitative distinction between -VCC.C- and -VCC- is attested, for example, in Qft (Upper Egypt; Nishio 1994: 41). For Egyptian dialects, that the incidence of i-deletion after medial geminates correlates with the possibility of initial clusters and with epenthesis of the -CCC- → -CIC- type was noted by Behnstedt and Woidich (1985: 74).

Shortening of nonfinal CVVC- before word-level endings, and in the output of word-level syncope, occurs only in CV-dialects: /baab-ha/ baabha 'her door', /saqṭib-ā saḥb-i 'my friend', /saqṭib-ā saḥb-ā 'girlfriend' (column F).

"Cyclic" effects, such as the contrast between ḫimnna 'we understood' and ḫimnna 'he understood us' (Brame 1974, Kenstowicz 1981, 1996), are attested only in VC-dialects.

Opaque epenthesis/stress interactions (such as the antepenultimate stress of /ẖim-nål/ ḫimmnā 'our understanding') are attested only in VC-dialects. Only here is the epenthetic vowel invisible to lexical processes like stress and vowel shortening. In CV-dialects, epenthetic vowels are always visible to lexical processes and get stressed under the same conditions as regular vowels; for example, Cairene /bint-nā/ biṭīna 'our daughter', like maktūba.

Previous theoretical work on two sets of instances by epenthesis and Levantine dialects. Dis (1981) in the context of are assigned to onsets in metathesis. On this approach, linked to syllable nodes; A rather different app approach developed by Farwaneh to right and that VC-dia noted that this process according to align.

Research on the see noted that in Levantine predictable and proposes Kenstowicz (1981, 1982) who worked out an OP and also proposed a tri-stress.

In the analysis to be differ in whether they to the prosodic word,1) Formally, semisyllables moras to be licensed by the form of syllables an section can be reconstituted: (1) C-dialects vocally, (2) VC-dialects a is undominated postlex morphological level (Licenser-μ) is is superior to the direct and eliminates the O/C and opacity phenomena.

Such theoretical into two aspects. First, it pro ering and, in particular, it is incompatible with constraint systems for v over must interact in stutes support for a c Morphology (LPM).
VC-dialects and are often 

\[ \text{itita}(h) \]

y in VC-dialects. V-dialects always under the pre- 

ident Broselow have ephenthesis of the VC-dialects. For instance, Bowles (1985: maps 59, 66), for interpreting a trisyllabic followed by luster (\(-\text{CIC-}\) 

1 VC-dialects. For example, (\(\text{ka} unfair\)), C- and C-dialects talk to someone, they teach'. All Iru (or \text{yikillimu}) is attested, for instance, in the possi- 

\(-\text{CIC-}\) type was understood, and in dialects: \text{baab-ha/}

\text{thb-a 'girlfriend' multistressed in VC-dialects, like stress and always visible to ions as regular, like \text{makiba.}}

Previous theoretical literature on these phenomena has concentrated primarily on two sets of issues: the divergent ways of resolving consonant clusters by epenthesis and the problems of cyclicity and opacity raised by the Levantine dialects. Discussion of the first question was launched by Selkirk (1981) in the context of syllable theory with a proposal that stray consonants are assigned to onsets in CV-dialects and to rhymes in VC-dialects. In a variation on this approach, Broseloe (1992) argued that stray consonants are linked to syllable nodes directly in CV-dialects and via moras in VC-dialects. A rather different approach was initiated by Ito (1986, 1989) and further developed by Farwaneh (1995). They suggested that CV-dialects syllabify left to right and that VC-dialects syllabify right to left. Mester and Padgett (1994) noted that this processual formulation can be translated into constraint-based terms by means of alignment constraints. Research on the second complex of issues began with Brrame (1974), who noted that in Levantine dialects the blocking of syncope is systematically predictable and proposed a cyclic analysis to explain it. It was continued by Kenstowicz (1981, 1983, 1986) and others and recently by Kager (1999), who worked out an Optimality Theory (OT) account using O/O constraints and also proposed a transferential analysis for the problem of opaque stress.

In the analysis to be developed in this chapter I claim that the dialects differ in whether they license unsyllabifiable consonants by moras adjoined to the prosodic word. I will refer to consonants so licensed as semisyllables. Formally, semisyllables arise where a constraint LICENSE-\(u\), which requires all moras to be licensed by syllables, is outranked by markedness constraints on the form of syllables and feet. The syllabic typology introduced earlier in this section can be reconstructed in terms of the level at which semisyllables are licensed: (1) C-dialects allow semisyllables both at the word level and postlexically; (2) VC-dialects allow semisyllables only at the word level (LICENSE-\(u\) is undominated postlexically), and (3) CV-dialects allow no semisyllables at any level (LICENSE-\(u\) is undominated everywhere). I show that this analysis is superior to the directionality/alignment approach to Arabic syllabification and eliminates the O/O constraints that have been proposed for the cyclicity and opacity phenomena.

Such theoretical interest as the analysis developed here lies primarily in two aspects. First, it provides evidence for the violable character of Strict Layering and, in particular, for moras that are unaffiliated with syllables. Second, it is incompatible with fully parallel OT, because it crucially requires distinct constraint systems for word phonology and sentence phonology, which moreover must interact in serial fashion. Thus, if the analysis is correct, it constitutes support for a constraint-based version of Lexical Phonology and Morphology (LPM).
The main features of the specific model of constraint-based LPM that I will assume are the following:

- Stems, words, and sentences are characterized by distinct constraint systems.
- These constraint systems are serially related.
- Morphology and phonology are cyclically interleaved in each domain.
- I/O constraints are the only type of correspondence constraint.

I refer to this model as stratal OT. In Kiparsky (to appear) I argue that stratal OT does a better job of explaining morphology/phonology interactions, and opaque constraint interactions, than parallel OT with an enriched correspondence theory. Trading in sympathy constraints, O/O constraints, and Paradigm Uniformity constraints for the intrinsic seriality of domains improves descriptively adequacy and lead to gains in learnability, naturalness, and typological restrictiveness. Just as LPM solves OT's problems with synchronic analogy and opaque constraint interactions, so OT helps LPM complete its synchronic program of modeling the lexicon and the morphology/phonology interface and its diachronic program of providing the basis for a theory of analogical change. In the cited work I support these claims with analyses of the principal benchmark phonological systems of the recent theoretical literature, as well as several that are new to it.

An important strand of evidence for stratal OT is that it helps realize the typological goals of OT phonology. Crucially, an important site of cross-linguistic and dialectal variation is whether a markedness constraint is active in stems, in words, or postlexically, which determines its domain and interaction with other constraints.

It is sometimes claimed that parallel OT is more restrictive than serial models. That is true only if correspondence theory is restricted to I/O constraints. When augmented with sympathy constraints, O/O constraints, and/or with Paradigm Uniformity constraints, parallel OT is not more restrictive than stratal OT. It is just different, and the differences are uniformly to its disadvantage.

While this chapter assumes stratal OT, and its results add a measure of support to that framework, its main purpose is not to compare it with parallel OT or with rule-based serial theories. It is primarily about syllable theory and the syllable structure of Arabic vernaculars.

6.2. Semisyllables

Generalizing the Exhaustive Syllabification principle of Selkirk 1981, the Proodic Licensing principle formulated in Ito 1986, 1989 requires that every syllable segment be assigned to a metric requirement. Stricter metrical element be in the next higher category (1986: 7):

(2) a. A given nonterminal of one or more
b. A unit of a given superordinate

In the OT perspective of Prosodic Licensing efficiently violations of Prosodic hierarchy constraints. Such violations are as I can tell. But St Arabic

I assume the proposal to a class of subconst the prosodic hierarchy in Arabic. Of particular interest and License-Segment, with a syllable and the constraints whose vernaculars.

The claim that Stric roots even in pre-OT w which violate Prosodic other types of prosodic been motivated by Ito: Hing. For metrical construct in (4).

(4) \[ \begin{array}{c}
\sigma \\
\mu \\
\end{array} \]

\[ \begin{array}{c}
\mu \\
\end{array} \]
segment be assigned to a higher-level prosodic constituent. A stronger licensing requirement, *Strict Layering*, requires that every non-highest prosodic or metrical element be in its entirety a constituent of an element belonging to the next higher category on the prosodic hierarchy (Nespor and Vogel 1986: 7):

(3) a. A given non-terminal unit of the prosodic hierarchy, \( X^* \), is composed of one or more units of the immediately lower category, \( X^{*-1} \).

b. A unit of a given level of the hierarchy is exhaustively contained in the superordinate unit of which it is a part.

In the OT perspective, higher-ranked constraints could force violations of Prosodic Licensing and *Strict Layering*. "Floating" elements are presumably violations of Prosodic Licensing, occurring when it is dominated both by markedness constraints (syllable structure, etc.) and by faithfulness constraints. Such violations of Prosodic Licensing do not occur in Arabic, as far as I can tell. But *Strict Layering* is, under certain conditions, violable in Arabic.

I assume the proposal of Selkirk (1995) that *Strict Layering* corresponds to a class of sub-constraints that regulate the affiliation of elements in the prosodic hierarchy and that are dominated but nonetheless visible in Arabic. Of particular interest here will be the two constraints LICENSE-\( \mu \) and LICENSE-\( \sigma \), which respectively require that a mora be affiliated with a syllable and that a segment be affiliated with a syllable. These are the constraints whose ranking determines the syllabic typology of Arabic vernaculars.

The claim that *Strict Layering* constraints are violable has respectable roots even in pre-OT work. Aside from extrametrical and floating elements, which violate Prosodic Licensing and a fortiori *Strict Layering*, several other types of prosodic representations that violate the latter constraint have been motivated by Ito and Mester (1992) under the heading of *weak layering*. For metrical constituency, Ito and Mester propose structures of the form in (4).

\[
\begin{align*}
&\sigma \\
&\mu \\
\end{align*}
\]
Structures of the form in (5), required for cliticization and compounding, respectively, also violate Strict Layering.

(5) a. 
   \[ \begin{array}{c}
   \ominus \\
   \downarrow \\
   \ominus \\
   X \quad \text{clitic}
   \end{array} \]

b. 
   \[ \begin{array}{c}
   \ominus \\
   \downarrow \\
   \ominus \\
   \ominus \\
   X \quad Y
   \end{array} \]

Ito and Mester claim that Strict Layering does hold between moras and syllables and formulate a principle ("Mora Confinement") that states that a mora can only be licensed by a syllable. This would be an unexpected restriction on their otherwise general hypothesis. I shall argue that the Weak Layering hypothesis holds in complete generality.

Violations of Weak Layering for moras require no novel constraints. The possibility is already inherent in the uncontroversial constraints of standard syllable theory. Suppose that FOOT-BINARITY, LICENSE-SEG (which requires that a segment be licensed by a syllable or mora), and the relevant MAX and DEP constraints all dominate the requirement that a mora be affiliated with a syllable (LICENSE-m). OT principles require, other things being equal, the representation that constitutes the minimal violation of the constraints. To avoid gratuitously violating Prosodic Licensing, a mora that cannot be affiliated with a syllable should be affiliated with the lowest possible superordinate prosodic category. In Arabic, affiliation of an unsyllabified mora with the next higher category foot would violate the otherwise undominated constraints on foot size. I will therefore assume that it is affiliated with the prosodic word, which is not subject to any size constraints (or at least not to any size constraints that would bar this affiliation).

Previous evidence for moraic licensing includes Hyman (1985) for Gokana, Zec (1988) for Bulgarian, and Buckley (1994). 6.4 for Kashaya. Moras that are prosodically licensed by adjunction to a superordinate prosodic category offer a way to accommodate what an older phonetic tradition originating with Sievers (1901) has called "semisyllables," or "minor syllables" (Sievers's "Nebensilben"). I shall adopt the term semisyllable here to denote such an unsyllabified mora, without meaning to imply that all the things that have been called semisyllables are necessarily to be analyzed that way.

Semisyllables offer what is arguably the right representation for trimoraic trochees, including superheavy syllables and resolved trochees. Bye (1997) argues that Estonian third unsyllabified mora unsyllabified mora in, does not seem to be at

In a number of Slavic in positions where th example, Russian word pronunciation, etc. show: consonant + liquid sequi

(6) a. Russian: mglä
b. Czech: rty 'lips'
   'wool', včerä 'w

In some Mont-Khmer labile (a "hl-syllable") (1983), who proposes prepended headless syllable, in his terminology out by Shaw (1983) for I also argue that some se

A more minimalist chapter, is that they are would be regular syllab this analysis for Bella C tion or sonority.

I do not wish to claim fact, it is possible that Kammu, Svanteson's (1 syllables. Nontonal min possibility of an added or Tonal minor syllables a the or segment, which may be would be a semisyllable syllable.

One virtue of the in offered here is that it is less sonants. If ONSET is al License-m, then onset as semisyllables. Initial vowels tend various languages, of a k
argues that Estonian and Saami superheavy (overlong, Q3) syllables have a third unsyllabified mora after the bimoraic core. He actually suggests that the unsyllabified mora in such sesquisyllabic structures is freestanding, but that does not seem to be an essential feature of his analysis.

In a number of Slavic languages, nonsyllabic sonorants occur at word edges in positions where they violate the sonority sequencing constraint. For example, Russian words like in (6a) are monosyllabic (as evidence from stress, versification, etc. shows). In Czech, word-initial liquids are nonsyllabic, while consonant + liquid sequences are syllabic, as in (6b) (Rubach and Bootj 1990).

(6)  a. Russian: mglitch ‘mish’. rt’a ‘mouth’ (gen.) (one syllable)

In some Mon-Khmer languages, the canonical word structure is a sesquisyllable (a “1+1-syllable”), carefully documented for Kammu by Svantesson (1983), who proposes that sesquisyllables are regular syllables with a prepended headless syllable (syllable with an empty nucleus, a “minor syllable,” in his terminology). A degenerate syllable treatment has been worked out by Shaw (1993) for Bella Coola semisyllables; Cho and King (this volume) also argue that some semisyllables are nonmoraic syllables.

A more minimalist view of semisyllables, in line with the proposal of this chapter, is that they are unsyllabified moras. Sesquisyllabic structures, then, would be regular syllables with a prepended mora. Bagemihl (1991) suggests this analysis for Bella Coola, where semisyllables are not constrained by position or sonority.

I do not wish to claim that headless syllables are ruled out in principle; in fact, it is possible that they occur, in addition to unsyllabified moras, in Kammu. Svantesson (1983) distinguishes between nontonal and tonal semisyllables. Nontonal minor syllables contain just a single consonant, with the possibility of an added schwa in careful speech; for example, emnal ‘to sow’. Tonal minor syllables are of the form CC-, where the second consonant is either a sonorant (e.g., hrmal ‘soul’) or a reduplicated copy of the stem-final segment, which may be an obstruent (e.g., riww ‘bellows’). The former would be a semisyllable in my terms, the latter a degenerate (headless) full syllable.

One virtue of the interpretation of semisyllables as unsyllabified moras offered here is that it implies that they are not necessarily restricted to consonants. If Onset is also added to the set of constraints that dominate License-μ, then onsetless syllables will be avoided by treating onsetless vowels as semisyllables. Several researchers have indeed noticed that onsetless initial vowels tend to have a special prosodically defective status in various languages, of a kind that suggests that they are not syllables in their
own right at least at some level of representation. Matsak and Hyman (1990), in their study of Kinande reduplication, argue that there are only CV-syllables at the stem level and that unsyllabified moras join syllables at a later stratum (or postlexically). Downing (1998) pursues this analysis and relates it to the Onset constraint of Prince and Smolensky (1993). Further evidence for the degenerate status of onsetsless syllables has been presented by McCarthy and Prince (1993) and Odden (1995). In terms of the constraints assumed here, a parse as unsyllabified moras follows straightforwardly from the constraint ranking in which ONSET, MAX, and DEEP dominate LICENSE-μ.

The basic argument for treating semisyllables as unsyllabified moras is that it immediately explains their characteristic cross-linguistic properties:

1. Unstressed, toneless, or reduced tonal contrasts
2. Restricted segmental inventory
3. Can be less sonorous than syllable nuclei
4. Restricted shape (e.g., no onset, or no branching onset, no coda)
5. Sometimes restricted to peripheral position (typically word edges)
6. Prosodically invisible
7. Can be subject to minimum sonority requirement

Precisely these properties hold for the second (nonhead) mora of syllabic nuclei. On the assumption that they are properties of nonhead moras, the present proposal predicts them for semisyllables as well.

In the C- and VC-dialects, LICENSE-μ is ranked below the constraints that impose syllable and foot well-formedness, and below REDUCE (Kirchner 1996, McCarthy 1999), which requires minimizing the duration of light syllables on the scale a > lu > 0 (in practice, in nonfinal position, because a dominant alignment constraint preserves vowels at the right edge). This ranking will result in certain sesquisyllabic structures. For example, on /baab-ha/, /yi-kiib-ii/, /?akil/, and /silab/, it imposes the syllabification shown in (8) in the C-dialects, and, at the word level, in the VC-dialects as well.\(^7\)

(8) Semisyllables in VC- and C-dialects

---

SYLLABLES AND MORAS

Postlexically, VC-diale under certain (parily (yi.ki)n, bu, (la)k, → from deletion, for exar

The CV-dialects, w constraint, the syllabi postlexical phonology.

(9) In CV-dialects, m

Crucially, moras that toward syllable weight labels permit what loc regular syllables with a

Additional evidence for prosodic minimalit lengthening in the dial.

87\) Like most dialect syllabic lexical words single mora because th

nands to be weightless accommodated to the r

the root vowel is length

l (or u, in back contexts short (see 10b)). As i

optionally get a profl

prothetic vowel is unit word minimum. There (see 10b).

(10) a. lkal/ koll 'eat', \(l\) imperatives)

b. ikil 'write', \(k\)

c. l\(\text{kk}\)n - l\(\text{kk}\)n (*
Postlexically, VC-dialects epenthize a vowel before moraic consonants under certain (partly variable) conditions, for example, \((yi)k_t\_bu \rightarrow (yi)k_t\_bu\), \((\dot{a}k)\_n \rightarrow (\dot{a}k)\_n\), and, optionally, in initial clusters that result from deletion, for example, \(s\_\_la\_\_n \rightarrow (li)\_\_la\_\_n\).

The CV-dialects, where avoidance of semisyllables is a high-priority constraint, the syllabification in (9) results instead and is retained in the postlexical phonology.

(9) In CV-dialects, moras must be affiliated with syllables.

Crucially, moras that are not affiliated with syllables or feet do not count toward syllable weight or foot size. Therefore, all dialects that allow semisyllables permit what look like superheavy syllables, that is, on our analysis regular syllables with an adjoined semisyllable.

Additional evidence that unsyllabified and unfooted moras do not count for prosodic minimality is furnished by the relationship of epenthesis and lengthening in the dialect of Đe: izzOr spoken in Syria (Jastrow 1978: 79–80, 87). Like most dialects of Arabic, Đer izzOr categorically prohibits monosyllabic lexical words of the form (C)CVC. Such words would have only a single mora because the undominated constraint Final = C forces final consonants to be weightless. When the morphology forms such words, they are accommodated to the minimum word requirement in one of two ways. Either the root vowel is lengthened (see (10a)), or alternatively, a prothetic stressed ı (or ü, in back contexts) may be added, in which case the root vowel remains short (see (10b)). As in other dialects of this type, all initial clusters may optionally get a prothetic vowel. Because epenthesis is postlexical, the prothetic vowel is unstressable and does not count toward satisfaction of the word minimum. Therefore, lengthening takes place regardless of prothesis (see (10c)).

(10a) a. /ka/ıl ‘kat’, /kt\ıb/ ‘kib’ ‘write’, /rk\ıb/ ‘rkib’ ‘climb’, /vr/nɐ ‘hit’ (2sg.m. imperatives)
b. /k\ıb ‘write’, /rk\ıb ‘climb’, /vr/nɐ ‘hit’
c. /sk\ı – /sk\ın (‘isk\ın’ ‘what’
Thus, on the one hand, /ktib/ 'write!' has three possible pronunciations in the Der izZor dialect: kteb, ikteb (lexical lengthening with optional postlexical epenthesis), and ikteb (lexical epenthesis). Impossible, on the other hand, are the following pronunciations: *kteb (too short), *ikteb (with postlexical epenthesis the word still does not satisfy the lexical word minimum; with lexical epenthesis the stress is misplaced), *ikteb (if postlexical epenthesis, it has the wrong stress; if lexical epenthesis, the vowel lengthening is unnecessary, constituting a gratuitous Dep violation).

Jastrow (1978) insightfully explains these and other data on the basis of a distinction between phonemic and phonetic epenthesis, which in this case is for practical purposes equivalent to our distinction between lexical and postlexical epenthesis. A similar idea figures in other traditional and structural grammars of Arabic as well.

6.3. Overview of the Analysis

The syllable structure of the dialects differs in the ranking of License-\(\mu\) in the word-level phonology. In the VC- and C-dialects, it is outranked by a number of faithfulness constraints (of both the Max and Dep type), by the markedness constraints Ft-Bin, License-C, and by Reduce (which minimizes the number of light syllables, specifically, because of dominating Align and Max constraints, of nonfinal light syllables with high vowels); in CV-dialects, it dominates them. The tableaux in (11) and (12) show the basic idea in schematic form.

Because of the low ranking the long vowel in /baabth/ /yiktibu/, in each case for dialects, however, License causing epenthesis of an \(\mu\) in the language (yi\(\kappa\)).\(\nu\).\(\mu\) rather than after the semi-minimal modification the in line with the language (\(\nu\).\(\kappa\).\(\nu\).\(\mu\)), not (*yi\(\kappa\)).\(\nu\).\(i\).

Nonfinal CVVC syllables in all VC- and C-dialects, a semisyllable (\(\nu\).\(\kappa\)).\(\nu\).h high at the word level, phonological constraints, eliminate CVVC syllables \(\nu\).\(\kappa\).h. Some do it by \(\mu\) dialects accommodate subtypes of CV-dialects a Dep-\(\mu\), Ft-Bin, and Max-.

Phrase-initial onset \(\nu\) for example, \(\kappa\).\(\nu\)\(\kappa\).\(\nu\)\(\mu\) - (i) CV-dialects. Syncope is al consonant can be licensed nates is similar: Syrian (i)
Because of the low ranking of LICENSE-μ in VC- and C-dialects, they preserve the long vowel in /baabha/ and delete the syllable final -i- in /thimma/ and /yik-tiib/, in each case forming an unsyllabified moraic consonant. In the VC-dialects, however, LICENSE-μ is promoted in the postlexical phonology, causing epenthesis of an unmarked vowel (i, or a, in those dialects that have it) before semisyllables. In the VC-dialects, the vowel is epenthized before rather than after the semisyllable because of prosodic faithfulness; it is the minimal modification that brings the word-level moraic (semisyllabic) parse into line with the language's surface syllable canon. Thus, (yik)t_u, bu → (yik)t_u, bu, dot ("yik".t_u, bu).

Nonfinal CVVC syllables that arise in the word-level phonology surface in all VC- and C-dialects, because these dialects license the third mora -C as a semisyllable ("baab", "baabha" ‘her door’). In the CV-dialects, LICENSE-μ ranks high at the word level, forcing violations of faithfulness constraints and phonological constraints, depending on their ranking. Most CV-dialects eliminate CVVC syllables by shortening the vowel, for example, /baab-ha/ → /baabha/. Some do it by epenthesis (baabha), and some Middle Egyptian dialects accommodate superheavy syllables unchanged (baabha). These three subtypes of CV-dialects arise by variation in ranking between the constraints DEP-μ, F1-BIN, and MAX-μ.

Phrase-initial onset CC-clusters are allowed only in VC- and C-dialects, for example, /s_t(aa)b_a/ → /l(is).t(aa)b_a/ ‘plowshare, weapon’, versus /s_t(laa)b_a/ in CV-dialects. Syncope is allowed to create initial clusters only where their first consonant can be licensed as a semisyllable. The distribution of initial geminates is similar: Syrian (Mitchell 1993: 92ff) /n-midd/ mmidd ‘we extend’,
/w-t-afaːl/ tafaː ‘he agreed’ (infixed /t/). Moroccan /ṭṭa-kṭoːb/ naktab ‘it was written’. The semisyllabic analysis of initial geminates is shown in (13).

(13) \[
\begin{array}{c}
\omega \\
\phi \\
\sigma \\
\sigma \\
\mu \\
\mu \\
\mu \\
\mu \\
\mu \\
\mu \\
\mu \\
\end{array}
\]

In addition to the dialectal distribution of geminates, this representation accounts for their phonological properties. They are bimoraic, hence true geminates, as shown by the fact that they commonly arise by assimilation. While the first member is moraic, it is nonsyllabic, hence invisible to stress, and does not satisfy word minimality constraints. Because final consonants and the initial consonant of a geminate are both semisyllables, a hypothetical *mm⁴nt would have a monomoraic foot; therefore, it is not a possible word.

Medial -CCC- clusters can be parsed in VC- and C-dialects by making the middle consonant a semisyllable; for example, /gil-t-la/ → (word level) (gil),t,la → (postlexical) (gil,t,la) in VC-dialects, but /ṭul-t-lu/ → (ṭul)(ṭ,lu) in CV-dialects.


In CV-dialects, word-level syncope is blocked after clusters, because they do not allow the semisyllable needed to parse its output in conformity with Fr-Brs, hence /yiktib-u/ → (yik)t, bu. C-dialects simply have deletion without postlexical ephenthesis. My analysis requires, and is in a sense a consequence of, the intrinsic serialism of levels in LPM. Any evidence for it will thus further support serial OT over parallel OT. The analysis of metathesis crucially depends on the serial relation between the word-level and postlexical phonology. The two constraint systems induce an intrinsic ordering /yikṭibu/ → yik-t, bu → (yik)t, bu. This is also the key to our solution to the opacity and cyclicity problems.

High vowels delete after geminate consonants only in VC- and C-dialects /y-kalim-u/ ṣ(i)kal(Im)mu ‘they talk to someone’ versus yikalimu in CV-dialects. Only in the former can the resulting superheavy syllable be prosodically licensed.

The correlation of medial syncope, “metathesis,” and retention of super-
heavy syllables follow because deletion of light vowels and retention of super-heavy syllables are prosodically licensed the same way; in VC-dialects they can apply more widely because semisyllables are allowed. Several intermediate forms of VC- and CV-dialects are also attested, corresponding to the possibility of ranking LICENCE-µ at several points among the other constraints.

Phrase-final -CC clusters that violate sonority sequencing occur only in CV- and C-dialects, but their status is quite different in the two types. In VC- and C-dialects, the second C is licensed as a semisyllable at the word level, and in VC-dialects an epenthetic vowel is inserted before it, at least if the cluster violates sonority sequencing. For example, the lexical representation of /akl/ is /aklā/ in VC-dialects and in C-dialects: the semisyllable is broken up by postlexical epenthesis in VC-dialects (lakil). In CV-dialects, in contrast, the second consonant is parsed in the word phonology as a nonmonomorphemic consonant (an "extrametrical" consonant adjoined to the prosodic word), namely /akā/: therefore, epenthesis is inapplicable.

Desonorization of word-final consonants seems to be a trait of CV-dialects. It characteristically involves devoicing and glottalization, sometimes near-deletion. In northern Yemen, the Western dialects break up final -CC clusters of rising sonority with an epenthetic vowel, a VC-dialect trait (e.g., /lism/ → (li)sim 'name', /lībn/ → (li)bin 'son'). Those of the eastern part seem to be of the CV-type, and lack epenthesis in -CC clusters. In these dialects, final consonants, including sonorants, are glottalized after a long vowel and devoiced in clusters: ÿ/Ç, -VCÇ. Jastrow (1989: 110) reports similar desonorization phenomena for southern Yemen, which also retain -CC. Epenthesis and desonorization thus seem to be in complementary distribution (Behnstedt 1985: 14, 48, 88). In some Egyptian dialects, again of the CV-type, word-final consonants are devoiced and an intrusive glottal is heard; for example, /-al/ → [aʔ], /-ooz/ → [ɔʔ] (Behnstedt and Woidich 1985, maps 41–43). This distribution can be understood if we take desonorization to be a process that applies to nonmonomorphemic consonants, in terms of the syllable structure of VC- and CV-dialects proposed in (8) and (9). Thus, the boxed consonant in (14b) undergoes desonorization.

```
14) a. VC-dialects
```

```
   w
  / \  \\
   \ /  \\
   e a k
```

Paul Kiparsky

*ktōb / tkōtōb 'it was shown in (13).*
That moraic consonants should tend to remain more sonorous than non-moraic consonants makes good phonological sense, since moraic elements are known to be subject to sonority restrictions (Zec 1988).

Cyclic blocking of syncope, typified by the ʃhinna: ʃhinna contrast, is restricted to a group of VC-dialects. All Levantine dialects are said to have it (Mitchell 1993:156). To my knowledge it has not been reported for any CV-dialect, and this is predicted, for in CV-dialects, deletion cannot give rise to forms like ʃhin-na ‘we understood’ in the first place, so the question of blocking that deletion by lexically assigned stress does not arise. In C-dialects, REDUCE outranks MAX-V, so that stem-level stresses do not protect vowels from deleting at the word level; thus, these cyclic effects cannot surface. The tableau in (15) shows the word-level (w.l.) phonology of VC-dialects; the corresponding tableau in (16) is for CV-dialects.

(16)

<table>
<thead>
<tr>
<th>CV-dialects</th>
<th>ʃhin-na ‘we understood’ (w.l.) phonology</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Input: [ʃi.him] ‘be understood’ (unsuffixed); no change</td>
<td></td>
</tr>
<tr>
<td>b. Input: [ʃi.(him).na] ‘we understood’ (stem-level suffix)</td>
<td></td>
</tr>
<tr>
<td>i. ʃi.(him).na</td>
<td>***</td>
</tr>
<tr>
<td>ii. ʃi.(him).na</td>
<td>*</td>
</tr>
<tr>
<td>iii. (him).na</td>
<td>***</td>
</tr>
<tr>
<td>c. Input: [ʃi.him]-na ‘he understood us’ (w.l. suffix)</td>
<td></td>
</tr>
<tr>
<td>i. ʃi.(him).na</td>
<td>***</td>
</tr>
<tr>
<td>ii. ʃi.(him).na</td>
<td>*</td>
</tr>
<tr>
<td>iii. (him).na</td>
<td>***</td>
</tr>
</tbody>
</table>

Invisibility of openthi (shortening) is restricted consonants are licensed at the word level.12 The unviolable requirement ʃhim-na → (w.l.) ʃh Stray consonants cannot receive an necessarily visible to Cairene (Tul)(fu) ‘/lye’.

In sum, apparent surt lexical stress assignmen ʃhim-na is disyllabic to. The assumption that lexical phonology of V of stress and epenthesis.

The licensing of constraints. To derive hypothetical forms will some constraint prohibit consequence: the two n cannot be licensed as making the closed syll: -il- cannot be parsed outlessless) as one. The its consonants to be pr
sonorous than non

sonorous than non-
moraic elements are

zioni contrast, is
cities are said to have
ported for any CV-
cannot give rise to
a question of block-
ris. In C-dialects,
not protect vowels
cts cannot surface.
nts of VC-dialects; the

<table>
<thead>
<tr>
<th>CV-dialects</th>
<th>*C</th>
<th>MAX-[V]</th>
<th>LICENSE-¿</th>
<th>REDUCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Input: [fi.¿m] 'he understood' (unsuffixed): no change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Input: [i.¿(him)na] 'we understood' (stem-level suffix)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. (\text{ff}) (him).na</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ii. fi.(him).na</td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iii. (him).na</td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Input: [fi.¿m]-na 'he understood us' (w.l. suffix)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. (\text{ff}) (him).na</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ii. fi.(him).na</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td></td>
</tr>
<tr>
<td>iii. (him).na</td>
<td>*</td>
<td>*</td>
<td>**</td>
<td></td>
</tr>
</tbody>
</table>

Invisibility of epenthetic vowels to lexical constraints (such as stress and shortening) is restricted to VC- and C-dialects, because in these dialects stray consonants are licensed as semisyllables, so epenthetic vowels are not present at the word level. This results in surface violations of the otherwise inviolable requirement that the last nonfinal foot is stressed; for example, /fi:m-na/ \(\rightarrow\) (w.l.) /fîhm\(_m\)\(_k\)/na \(\rightarrow\) (postlex.) fîhim-na 'our understanding'. Stray consonants cannot be licensed in CV-dialects (except word-finally), so they must receive an epenthetic vowel in the lexical phonology, which is necessarily visible to stress like any other lexical vowel. For example, Canarese (?u)/(i/lu) 'I/you (m.sg.) said to him' is stressed like (mak)/(i/da).

In sum, apparent surface exceptions to stress arise when the conditions for lexical stress assignment are masked by postlexical epenthesis. A word like fîhim-na is disyllabic for purposes of the word phonology, including stress. The assumption that unsyllabifiable consonants are semisyllables in the lexical phonology of VC-dialects explains the opaque constraint interaction of stress and epenthesis.

The licensing of semisyllables in VC-dialects is subject to certain constraints. To derive ka(âb)\(_k\)/\(_k\) rather than *(kâta)b\(_k\)/\(_k\) and to exclude hypothetical forms with complex consonant clusters, we must assume that some constraint prohibits two adjacent semisyllables. This has a further consequence: the two middle consonants of a medial four-consonant cluster cannot be licensed as consecutive semisyllables. Epenthesis is then lexical, making the closed syllable visible to stress. In an input like /ki\(n\)b-t\(l\)-ha\(_l\)/, -\(tl\)- cannot be parsed as two semisyllables, nor (since semisyllables are onsetless) as one. The cluster must therefore get a full syllable for all its consonants to be parsed. A full syllable requires a vocalic nucleus and
therefore gets an epenthetic vowel, forming a closed syllable at the word level. Contrast:

(17) a. /kitab-I-ha/ → (w.l.) kitább, ha → (postlex.) kitább,ha 'he wrote to her'
b. /kitab-il-ha/ → (w.l.) kitabíllha 'I wrote to her'

In the following sections I motivate in more detail some aspects of the analysis just sketched out.

6.4. Initial Geminates

Mitchell's (1993) phonetic description of initial geminates in Arabic states clearly that they share with medial geminates the phonetic characteristics of length and tenseness: "All types of gemination reveal not only an increase of duration over non-gemination but also greater muscular tension and pulmonary pressure, a more extensive spread of tongue-palate contact, increased loudness of adjoining vowels and 'indisiveness' of on- and off-glides (especially in the case of plosive consonants), as relevant phonetic characteristics. ... An epenthetic vowel may be heard initially in most cases of initial gemination but it is never essential and is better omitted in the [Moroccan, Iraqi, and Levantine] vernaculars." (Mitchell 1993: 92). Owens (1984: 26) similarly states that gemination in eastern Libyan "may have a realization as glottalization or an increased length on an initial consonant, the latter realization common before voiceless consonants and nasals."

Initial geminates constitute fairly direct evidence for unsyllabified moras. Given that assimilation is the spread of a melody over timing slots or syllabic positions, an assimilated geminate must be affiliated with two such positions. In a medial geminate, these positions are the coda of one syllable and the onset of the next, or, in the most popular version of the moraic theory, the second mora of one syllable and the syllable node of the next. Onset geminates must have some extra syllabic position, and under standard assumptions, this can only be a mora. (It cannot constitute a syllable of its own, since the geminate consonant does not function as a syllable peak.) This mora must be affiliated with the foot or the word, presumably depending on whether it adds weight. Thus, moraic theory requires representations like (18a) for medial geminates, as in a hypothetical word atta, and (18b) or (18c) for initial geminates, as in not.

(18) a. σ
   \[ \sigma \]
   \[ μ \]
   \[ \mu \]
   \[ a \]
   \[ t \]

If we adopt this representation, we will immediately explain why initial clusters. In Arabic, it seems those VC-dialects that also categorically excluded.

(19) a. CV-dialects: Egyp
   \[ fihtml 'I wrote', fihtml 'I \]
   b. VC-dialects: Lev
   \[ shard, claab - ?. \]
   (Syrian; Grotfeld
   c. C-dialects: kūbaIr
   \[ šīmōl camel mea \]
   62-65, kūbaIr 'I \]
   (Mzāb), kūbaIr 'I \]

Many initial geminates infixes to stem-initial cons seems to mark passive...
syllable at the word

*ilha* 'he wrote to her'

some aspects of the

tes in Arabic states
the characteristics of
only an increase of
ar tension and pul-
te contact, increased
nd off-glides (espe-
etic characteristics.
osi: cases of initial
l in the [Moroccan, 
> Owens (1984: 26)
ave a realization as
ant, the latter reali-
insyllabified moras.
ing slots or syllabic
two such positions.
le syllable and the
oma theory, the
xt. Onset gemi-
stand assump-
able of its own, since
ik.) This mora must
iding on whether it
ons like (18a) for
or (18c) for initial

If we adopt this representation of initial geminates for Arabic, we can immediately explain why their distribution correlates with the distribution of initial clusters. In Arabic, initial geminates are restricted to C-dialects and to those VC-dialects that allow initial clusters. In CV-dialects, initial clusters are categorically excluded.

(19) a. CV-dialects: Egyptian *kitaab* 'book' (*ktaab*), *silāh* 'plowshare', *katāb* 'I wrote', *fūnūn* 'I understood'
b. VC-dialects: Levantine *ktaab* ~ *ktaab* 'book', *silāh* ~ *ilsilāh* 'plow-
share', *claab* ~ *isilaab* 'dogs' (Iraqi; Erwin 1963), *ktāb* 'I wrote'
(Syrian; Grotfeld 1980: 177), *fi:lūn* 'I understood'.
c. C-dialects: *ktāb* 'I wrote', *fūnh* 'freesq' ‘at the side of the road’, *ilhān*
*ilma* 'camel meat', *bīlīya* 'in the evening' (Moroccan; Mitchell 1993:
62-65), *ktāb* 'I wrote', *klaab* 'dogs' (Tunis), *nhag* 'he was burned'
(Mzāb), *ktb* 'I wrote' (Benghazi (Singer 1980: 255, 257, 260).

Many initial geminates arise by assimilation of consonantal prefixes or
fixes to stem-initial consonants ((20a–g)). In Maghrebi dialects, initial [t]–
seems to mark passive verbs ((20h–i)) and stative verbs ((20j)).
CV-dialects admit no initial consonant clusters, hence in particular no initial geminates, so prothetic [ʔ:] is obligatory in cases comparable to (20). Similarly, those VC-dialects that avoid word-initial clusters tend to have obligatory prothesis before initial geminates, such as Cyrenaican Bedouin.

Therefore, I take the parallelism between initial geminates and consonant clusters as support for the semisyllabic analysis.

Could the parallelism be accounted for without invoking mora theory? Owens’s (1980) structural analysis of eastern Libyan is one analysis that attempts to do that. He takes the first member of an initial geminate as the realization of a vowel “v”, which is the epenthetic vowel here written t (21a) is his representation of an initial geminate, and (21b) shows a variant articulation where the first member is replaced by a glottal.

(21) a. j j ‘skins’
   \[CC\]

b. ? t tikāllam ‘you m. speak’
   \[CC\]

To update something like Owens’s analysis, we would have to reintroduce CV slots or X-slots into phonological representations (Clements and Keyser 1983; Levin 1985), and that is just what Hume et al. (1957) advocate. With one crucial exception, their data on Leti gemination is very similar to the Arabic data discussed here, but they arrive at a different conclusion. In both languages, geminates pattern distributionally with consonant clusters (in Leti, geminates and clusters occur underlyingly only in word-initial position). They arise in medial as well as in initial position by assimilation of adjacent consonants and, unlike other two-consonant sequences, may not be broken up by epenthesis (for Arabic, see Abu-Salim 1980) — evidence that they have a single root node. And in the weight of the syllable.

Hume et al. nonetheless has to do with an opt prosodic prominence of certain concomitant vox and words with geminat downgrading process. His PROMINENCE constraint, v domain, and plausibly submerger of durational consistent with the mora (i.e., mora unattached to a syllable-initial syllable vowels in terms of weigh extrasyllabic mora were phonological phrase” (Ht). Actually, there is no di

guish carefully between a segment is long if it is syllabic. And let us say ti (and, of course, superheated consonant in (18b), is long with does not contribute syllable. The LENGTH TO F of segment length, exactly to long vowels and geminat.

A second objection that as in (18b), is that it viola the unsyllabified mora is s with a higher prosodic cate: objection does not seem v.

My proposal immediate initial geminates are not e be languages in which a perhaps at the foot level. F type (see also Davis 1990 minimum and an undominant: weightless. Because final t words like maa ‘behavior’ and *ba and *but do not occ of these otherwise forbid
SYLLABLES AND MORAS IN ARABIC

single root node. And in both languages, initial geminates do not count toward the weight of the syllable that follows.

Hume et al. nonetheless reject representations like (18b–c). The reason has to do with an optional postlexical process that “downgrades” the prosodic prominence of the first of two syntactically related words, with certain concomitant vowel reduction processes. Words with long vowels and words with geminates including initial geminates cannot undergo this downgrading process. Hume et al. capture this restriction with a LENGTH-TO-PRONUNCIATION constraint, which requires a long segment to be in a prominent domain, and plausibly suggest that its functional motivation is to prevent the merger of durational contrasts. They then state that this constraint is inconsistent with the moraic (semisyllabic) representation of geminates: "Since a mora unattached to a syllable node would not contribute to syllable weight, a geminate-initial syllable would again not be predicted to pattern with long vowels in terms of weight. Note that this same problem would arise if the extra syllabic mora were linked to a node higher than the syllable, e.g. foot, phonological phrase" (Hume et al. 1997: 393).

Actually, there is no difficulty with the moraic representation if we distinguish carefully between segment length and syllable weight. Let us say that a segment is long if it is affiliated with more than one timing slot (mora or syllable). And let us say that a syllable is heavy if it has more than one mora (and, of course, superheavy if it has more than two moras). Then the initial consonant in (18b), is long (bimoraic), but the first mora that it is affiliated with does not contribute to syllable weight, because it is not part of the syllable. The LENGTH-TO-PRONUNCIATION constraint can then be stated in terms of segment length, exactly as Hume et al. have it, and it will correctly refer to long vowels and geminates, including initial geminates.

A second objection that Hume et al. raise to the moraic representation, as in (18b), is that it violates Prosodic Licensing. But this is the case only if the unsyllabified mora is stray. Since the unsyllabified mora can be affiliated with a higher prosodic category, as they recognize in the passage quoted, this objection does not seem valid either.

My proposal immediately entails that there could exist languages in which initial geminates are not only long but weight bearing as well. These would be languages in which a stray mora is associated at the syllable level or perhaps at the foot level. Flart (1991) shows that Trukese is a case of just this type (see also Davis 1999). Like Arabic, Trukese has a bimoraic word minimum and an undominated constraint that requires final consonants to be weightless. Because final consonants do not contribute to syllable weight, words like maq ‘behavior’ and tiq ‘emotions’ are possible, whereas words like *baa and *ban do not occur. Yet words with initial geminates allow both of these otherwise forbidden rhyme types; for example, no ‘clam sp.', ffin
advice'. The conclusion is that in Trukese initial geminates contribute to prosodic weight and that the weight-bearing mora must be contained within the prosodic category in terms of which the minimality constraint is stated. Supposing this to be the foot, the structure of *to* in Trukese would be as shown in (18b).

A second, similar case would be Piro (Lin 1997). Lin works out a moraic licensing analysis of stray consonants, including initial geminates. Since Piro has compensatory lengthening effects, the semisyllables would have to be associated at the foot level. This appears to be compatible with Lin's analysis, though Lin opts for the alternative affiliation at the prosodic word level.\(^{15}\)

### 6.5. Sonority, Syllabicity, and Epanthesis

Nearly all CV- and C-dialects allow phrase-final -CC clusters.\(^{36}\)

\begin{enumerate}
\item a. CV-dialects: *kāḏir* 'I wrote', *ḥakl* 'food', *gabr* 'algebra' (phrase-finally), but *ḥakl mar* 'Egyptian food'
\item b. C-dialects: *kāḏir* 'I wrote', Moroccan *šāqī* 'I needed', *rābī* 'winning', *ṣām* 'body' (Mitchell 1993: 70ff., Cohen 1975: 74ff.)
\end{enumerate}

With regard to their treatment of phrase-final -CC clusters, VC-dialects fall into two areally discontinuous groups. The first group of dialects permits phrase-final -CC clusters only if they satisfy the Sonority Sequencing Principle. Most Levantine dialects are of this type. The typical VC-dialect pattern is illustrated by the examples in (23) from Tripoli (Kenstowicz and Abdul-Karim 1980).

\begin{enumerate}
\item b. /ḥilm/ ḥilm 'load' ḥilm-ɨ ḥilm-ak
\end{enumerate}

A second group of VC-dialects permits no phrase-final -CC clusters at all. For example, on the one hand Baghdad Christian Arabic (Abu-Haidar 1991) epanthetizes a into a final cluster irrespective of sonority, and there seems to be no evidence of an underlying contrast between /CVCaC/ and /CVCC/\(^{17}\)

\begin{enumerate}
\item a. kāḇb 'dog', bānāt 'girl', šār 'drink'
\item b. šām 'name', māṣr 'Egypt', qāmāl 'lice'
\end{enumerate}

In Baghdad Jewish Arabic, on the other hand, epanthesis is sensitive to the sonority profile of the final cluster, and /CVCaC/ contrasts lexically with /CVCC/ (Mansour 1991: 10ff.):\(^{18}\)

\begin{enumerate}
\item a. /bānt/ bānt 'daughter', /bānt-al bānta 'her daughter'
\item b. /bānt-al bānta 'she built it' (L)
\end{enumerate}

A number of writers: sis in -RC clusters of fa 80f., Haddad 1984).\(^{19}\)

(26) a. /bīn/ bīn - bi stand?', /dars/ a
b. /ḥilm/ ḥilm 'arec

A second point of va
In all CV-dialects, epanthe
ty vowels. In the ma
stress, both in that th
able they head is 'skip
example in Baghdad Je
34), and in the Anatoli

(27) a. kāḇ-ki 'your
b. āḵ-나 'our to

But there are dialects i
stress:

(28) āḵ-나 'our food'

In Baghdad Moslem A
variation with respect t
/gl-t-lā giłša - gišša 'i
hits him' - yudārbu (B
1993: 82, 194).

An epanthetic vow
the stem level or at th
sis of stem-level (unde
posibility (as suggeste
either case, we predict
between /CVCaC/ (or
true in general.\(^{20}\)

An interesting test t
studied by Johnson (11
specific classes of word
types of nouns in -CC t
is not of rising sonori
nonetheless counts as e
ples, such as dāres 'stud
and plurals of adjective
speaker, then, the mo

(26) a. /bin-t/ bin-t ‘daughter’, /fi:n-t/ fi:n-t ‘did you understand?’, /dars/ dars ‘lesson’
   b. /fi:n/ fi:n ‘dream’, /jibni/ jibni ‘cheese’

A second point of variation is whether the inserted vowel counts for stress. In all CV-dialects, eponhetic vowels behave with respect to stress like ordinary vowels. In the majority of VC-dialects, eponhetic vowels are invisible to stress, both in that they do not get stressed themselves and in that the syllable they head is “skipped” in the calculation of stress. This is the case, for example, in Baghdad Jewish and Christian Arabic (27a); Abu-Haidar 1991: 34), and in the Anatolian dialect of Mardin (27b); Jastrow 1980b: 144).

(27) a. /dã:m-ki/ ‘your dog’, /bõn-ka:m/ ‘your daughter’, /swm-na/ ‘our name’
   b. /swm-ka/ ‘our food’, /bõn-ki/ ‘your (f.) son’

But there are dialects in which the eponthetic vowel obligatorily counts for stress:

(28) /swm-ka/ ‘our food’, /bõn-ki/ ‘your (f.) son’ (Qartmin, Jastrow 1980b)


An eponthetic vowel that is stressable must be inserted lexically, either at the stem level or at the word level. We will assume the latter, though reanalysis of stem-level (underlying) representations should be kept in mind as a possibility (as suggested by Hamid [1984] for Sudanese nominal stems). In either case, we predict that such dialects will show no postlexical distinction between /CVCoC/ (or /CVGtC/) and /CVCC/. This prediction seems to be true in general. 21

An interesting test case appears in the Palestinian dialect of the speaker studied by Johnson (1979). Here eponhesion shows special behavior in two specific classes of words. The larger consists of some morphologically defined types of nouns in -CC that undergo regular eponhesion even when the cluster is not of rising sonority. For stress and postlexical phonology, the vowel nonetheless counts as eponhetic. These include infinitives and active participles, such as /ðãræ/ ‘studying’, versus the phonologically expected /ðãs/ ‘lesson’, and plurals of adjectives of color and defect, such as /zõr/ ‘blue’ (pl.). For this speaker, then, the moraic parsing of the final consonant in these specific
classes of words is fixed lexically, perhaps in observance of some templatic constraint; adjectives of color and defect notoriously show a rigid template and special morphophonological behavior in Arabic.

The second set is very small and contains obligatorily just the word kotob ‘book’ (and optionally also mâleq ‘king’). Here the vowel of the second syllable is invisible for purposes of word stress, like a normal epenthetic vowel (kôtobna ‘our book’, like fôrômna ‘our oven’), but it functions like an underlying vowel in the phrasal phonology (kotob el wâlad ‘the boy’s book’ versus fôr el wâlad). From the present point of view, this exceptional word must involve a lexical alternation between a bound allomorph /kotb/ in kôtobna and a free allomorph /kotob/ in kotob el wâlad.

The North African C-dialects would require special study, but a few inconclusive remarks may not be out of place here. According to phonetic descriptions, medial three-consonant clusters are either retained without epenthesis (yitbu, yiktbu) or get a furtive phonetic transitional vowel on one or the other side of the middle consonant: yik'tbu, yik'tbu (Fischer and Jastrow 1985: 65, Singer 1985: 255). The latter happens typically when the cluster contains a sonorant or guttural; for example, (Mauretanian) yaḍr'gu ‘they (will) ask’, yaḍ'hla ‘they (will) enter’, yar'dfu (Cohen 1963: 90).

When no epenthesis takes place, what is the syllable structure? One possibility is that consonants can be syllable peaks, at least on the phonetic level. Alternatively, the long consonant sequences of these dialects could be considered clusters rather than syllables (Mitchell 1993: 72). The representation suggested here raises a third possibility: that they are composed of moraic but unsyllabified consonants, that is, semi-syllables.

Auditorily, these dialects certainly have voiceless peaks. See Harrell 1962b, where various analyses including voiceless schwa and empty nuclei are considered. In songs, consonants (at least sonorants and voiceless fricatives) readily occupy beats. This could point to syllable status, but a semi-syllabic analysis cannot be excluded a priori.

My methodological starting point is that phonological rather than phonetic evidence is the key to the answer. Syllable structure, like stress, is a matter of prosodic organization and is not necessarily manifested directly in any single phonetic dimension. From comparative Slavic phonology (Liewehr 1967) it is clear that the difference between, say, a semi-syllabic r and a syllable r is not reliably identifiable from the phonetic record alone. The decisive considerations are whether it can or cannot be stressed, whether it can have an onset, whether it contributes to syllable and foot weight, and so on. Moreover, a segment can be nonsyllabic in the lexical phonology and syllabic in the postlexical phonology, as the n of English rhythm.

Phonological evidence of this kind is not easy to come by. However, there are indications that sonorants can be syllable nuclei in Maghrebi Arabic, at least on some level. For a word, as in Moroccan in this word were not empty nucleus, both pr such as Marrakhi mâkk

Conversely, it seems that obstruents are final obstruents cannot form level. If obstruents could and given these dialect coda as well, predicting to occur.

Cohen (1963: 128) de sonorants in Mauretanian

(29) a. ânفتح b. ânزراحøt nza

On the assumptions the described by Mitchell (it is a semi-syllable, thf, yoway Syllabificati be a syllable nucleus if Otherwise it is optional n, żaraḥøt). Under the between syllable and section, except when other parsing obligatory.

Cohen also cites the a reversal between und citation form is monosyllable, usually inserted, unlike ṣ̱ṯ. In the imperative (ṣ̱ṯ ‘write!’), but the all contexts, is monosyllabic form, we are in tro same time /ṣ̱ṯ/ → ṣ̱ṭ.

We can make some forms /kotb/ ‘he wrote’
The citation form can be ogy, derived by imposi constraint REDUCE, with perfect's initial C is join
least on some level. First, they can constitute the most sonorous segment of a word, as in Moroccan /drb/ [drb] ‘cul-de-sac’ (Mitchell 1993: 93). If the r in this word were not syllabic, then the word would have no syllable or an empty nucleus, both problematic options. Second, sonorants are stressable, such as Marrakshi mšıḥa (Mitchell 1993: 201).

Conversely, it seems that no word can consist entirely of obstruents and that obstruents apparently cannot be stressed, so we might conclude that obstruents cannot form syllable peaks in Maghrebi, at least at the phonetic level. If obstruents could be syllabic, they should be capable of having onsets, and given these dialects’ preference for closed syllables, they should have codas as well, predicting the possibility of words like *ktb, which do not seem to occur.

Cohen (1963: 128) describes the variation pattern in (29) for phrase-initial senerators in Mauretanian Arabic.

(29) a. anjara *njara 'I have been injured'
   b. anjaratı njara 'she has been injured'

On the assumptions that Cohen’s transcription an represents syllabic [n] as described by Mitchell (1995) and others, and that a in a sonority peak position is a semisyllable, the distribution in (29) could be understood in the following way. Syllabification is obligatory before -CV, because the nasal must be a syllable nucleus if it has to support a coda consonant, namely, njaratı. Otherwise it is optionally syllabified or remains a semisyllable (njaratı = njaratı). Under the assumptions stated, this would indicate variation between syllabic and semisyllabic parsing of sonorants in sonority peak position, except when other constraints on syllable structure make the syllabic parsing obligatory.

Cohen also cites the following remarkable data for Mauretanian, virtually a reversal between underlying and output forms. In the 3sg.m. perfect, the citation form is monosyllabic (ktrb ‘wrote’), but in context a prothetic ơ is usually inserted, unless the preceding word begins with a vowel, namely ơktob. In the imperative, the citation form is, on the contrary, disyllabic (ơktob ‘write!’), but the most frequent pronunciation in running speech, in all contexts, is monosyllabic ktrb. If the citation form is taken as the underlying form, we are in trouble, for how can we have /ktob/ → ơktob and at the same time /ktob/ → ktrb?

We can make some sense of this by positing the respective underlying forms Ơktob/ ‘he wrote’ and /kto/ ‘write!’ with lexical epenthesis to ơktob. The citation form can be identified with the output of the word-level phonology, derived by imposing on the underlying form the previously discussed constraint REDUCE, which bars light syllables. In the phrasal context, the perfect’s initial C is joined into a syllable with a preceding vowel. The imper-
ative is not subject to phrasal resyllabification, presumably because it initiates its own intonational group (as in English and in most languages) and only undergoes optional reduction to a monosyllable:

\[
\begin{align*}
\text{Perfect} & \quad /kətəb/ \quad ʷkətəb \quad \text{Cakətəb Vkətəb ‘he wrote’} \\
\text{Imperative} & \quad /ktəb/ \quad ʷktəb \quad ʷktəb \sim ʷkətəb \quad \text{‘write!’ (m.sg.)} \\
\end{align*}
\]

I tentatively conclude that sonorants in Maghrebi dialects may be syllabic if they are sonority peaks and that they are otherwise at least optionally retained as semisyllables.

6.6. Opaque Shortening of Medial CVVC

The subject agreement endings in Arabic are introduced at the stem level of the lexical phonology, while object endings are word level. Before consonant-initial subject endings, the long vowel of CVVC-syllables is obligatorily shortened in all Arabic dialects, regardless of whether the final cluster is subject to epenthesis; for example, Iraqi /qal-t/ giit, Egyptian /qal-tu/ gula‘ ‘I said’. So “CVVC syllables are categorically prohibited everywhere in the stem-level phonology. More generally, the Fr-Bnv constraint that prohibits feet exceeding the two-mora limit is undominated at the stem level.

The relationship between shortening before subject endings and epenthesis is opaque in all dialects (counterbleeding). For example, the 3sg form corresponding to saaf ‘he saw’, šif-na ‘we saw’ is šift or šaft; there are no dialects with forms like *šaaf-it ‘I saw’. This follows from the consonantal underlying form of the ending /-t/, if subject endings are assigned at the stem level in all dialects (as other facts of the phonology and morphology independently show). Epenthesis, in contrast, is active only in the postlexical constraint system or, in some dialects, at the word level, and therefore in either case intrinsically follows shortening.

With regard to the treatment of CVVC-syllables before object elisions and other word-level endings, dialects differ. VC- and C-dialects always retain the long vowel, or shorten it variably at the postlexical level, and CV-dialects show three different treatments: shortening, epenthesis, and retention of the long vowel:

\[
\begin{align*}
& (31) \quad /bāb-ha/ \quad \text{‘her door’} \\
& \begin{align*}
& \text{a. VC- and C-dialects: bābha} \\
& \text{b. CV-dialects} \\
& \quad \begin{align*}
& \text{i. Shortening: bābha (Cairo, Delta)} \\
& \text{ii. Epenthesis: bābhā (alFarāfira, adDīxila), bābhā (Mecca)} \\
& \text{iii. Retention of CVVC.: bābha (middle and northern Upper Egypt, ilBalahariyya)} \\
\end{align*}
\end{align*}
\]
In addition to object endings and possessive endings, a second class of word-level closed syllable shortening environments arises when medial syncope before vocalic word-level suffixes creates closed syllables, which are then shortened if long in the CV-dialects (Woidich 1980: 213, Abu-Mansour 1991):

(32) a. CV-dialects: kaatib-ī → kaṭaba ‘having written’ (L), saahib-ī → saḥba ‘female friend’

b. VC-dialects: kaatib-ī → kaṭaba, saahib-ī → saḥba

The dialectal distribution of this second type of shortening tallies very closely with that of the first, as we would expect.

Next consider dialects where superheavy syllables arising by morphological combination are eliminated by epenthesis, with preservation of the long vowel. They include two oasis dialects in Egypt (dialects (c-v) and (c-vii) in (1)), and Makkah, where the epenthetic vowel is a: /Sumr-ha/ → Sumrāha, /uff-t-nil/ → suf-tani (Abu-Mansour 1991: 141).

The third solution is found in Middle Egyptian CV-dialects. In these dialects, superheavy syllables occur, but the restrictions on consonant clusters are the same as in other dialects (ḥaṭṭa but yikību). A similar system is found in Makkah (Abu-Mansour 1991).

Typically, CVCC and CVVC syllables have a parallel status, and syllabification works to avoid both. Yet CVCC and CVVC do not behave in exactly the same way in all dialects; the former are more restricted. Also, syncope is allowed more readily after geminates than after other CC clusters, for shortening of long vowels and of geminates is a way of accommodating the syllable structure without incurring a melodic max-violation. This is an instance of a “lookahead” effect of the sort that OT makes scarce of.

The interplay of epenthesis and shortening in the resolution of overlength gives a measure of support for the LPM-OT approach to constraint interaction. Farwaneh (1995: 152) notes that monomorphic -CV endings bimorphic -C-V endings, and bimorphic -C-CV endings all behave differently. In the Levantine dialects described by Abu-Salim (1980) and Haddad (1984), only monomorphic -CV suffixes like those in (33b) allow the length of the preceding stem to surface.27

(33) a. Shortening before stem-level consonantal endings
   i. āṣaf /ṣaf/ ‘he saw’
   ii. /ṣaf-t/ šīt ‘I saw’
   iii. /ṣaf-nā/ šīfa ‘we saw’

b. No shortening before word-level -CV endings
   i. /beet-na/ bātina ‘our house’
   ii. /nnoś-ha/ nūšha ‘its (L) type’
   iii. /ṭī̇ntaaj-ha/ ṭīṭaajha ‘he needed it (i.)’
   iv. /jaab-ha/ ḫaṭba ‘he brought her’
c. Shortening before word-level -C-V endings
   i. /jaab-l-il/ jābilah 'he brought for me'
   ii. /jaab-l-ul/ jābiluh 'he brought for him'

d. Shortening before word-level -C-CV endings (with epenthesis)
   i. /jaab-l-ha/ jābiilha 'he brought for her'
   ii. /saaf-l-ha/ šāfoilha 'he saw for her'
   iii. /raah-l-ha/ rāhilha 'he went to her'

These data indicate that suffixation with -l triggers cyclic shortening of the resulting hyperlong (four-mora) syllable. This confirms that the word level can be cyclic and that bound bases can constitute domains of constraint evaluation.

Farwaneh’s South Palestinian dialect differs in having no shortening in the (33c) cases, such as jaahla (1995: 162). A third group of dialects keeps the vowel long before all word-level endings, including cases like (33c) and (33d); for example, Northern Iraqi Jewish Arabic baat-l-ul-me 'he sold them to him', jaah-al-kam 'he brought you (pl)' (Jastrow 1900: 59, 325).

A prediction of my analysis is that there should be no word-level closed syllable shortening before stressable epenthetic vowels. Because closed syllable shortening is lexical, it must be bled by lexical epenthesis, and it cannot be bled by postlexical epenthesis. The status of epenthesis as lexical or postlexical is independently determined by its interaction with word stress. Since stress is lexical, an epenthetic vowel that is stressable must be lexically inserted, and an epenthetic vowel that is unstressable must be postlexically inserted. So, closed syllable shortening should be bled by the insertion of stressable epenthetic vowels. This is confirmed by Soukline Syrian kaudilha 'he said to her' (Behastedt 1994: 107) and by Iraqi jaabilha (alternating with jaabilha) 'he brought to her' (Erwin 1965: 143, 41), with transparent constraint interaction (not *jaabilha).

Conversely, since unstressable (postlexically inserted) epenthetic vowels are invisible at the word level, they should not block lexical closed-syllable shortening. And this is what we find, for example, in Levantine jaab-l-ha → (word level) jābil, ha → (postlexical) jābilha. Thus the intrinsic serial relation of the levels enforces opaque (counterbleeding) constraint interaction.

In dialects with general pre-stress shortening, the root vowel is of course short even before a stressable epenthetic vowel. For example, all Egyptian dialects have pre-stress shortening (Fischer and Jastrow 1980: 213), evidently at the word level. Hence, Egyptian dialects that stress epenthetic vowels do shorten before those vowels, such as il-B’raa si jabilhum (Woidich 1973–1974: 365). Blanc (1953: 44, 75) explicitly gives this analysis for North Palestinian examples like ūmāar-l-kul/ ūmārikku 'your female donkey'. Similarly, the Benghazi Libyan jaabilha (Farwaneh 1965: 141) may really be jaabilha with the variable phonetic (1984: 30) for a closely

The prediction is that vowels only if they are from eastern Libyan in the output and apj -CVVC-C- in cases like (34) ma šif-nā-kā 'we

ma inšām-s 'not
daff 'has pushed'
gbli-ī-ha, from /g

Perhaps structures like semiconsyllables in a row

The dialect data are
diction that closed-syll
level or postlexical pr

tend to further suppor
is transparent within a

6.7. Summary

I have argued that unsyllabified moras (c)
These moras are licensed into the output for the Berber-type sy
epenthetic vowel is ins
lexical processes like at
the epenthetic vowel I
moraic character and t
In CV-dialects, a licen
syllables ranks higher,
consonants never becc
its place again dictate
differences between a
assumptions.

NOTES

1. I present a more exte

Effects and Opaque.
the variable phonetic shortening of unstressed vowels described by Owens (1984: 50) for a closely related Libyan dialect.29

The prediction is that dialects retain length before unstressable epenthetic vowels only if they tolerate hyperlong syllables. This is confirmed by data from eastern Libyan (Owens 1984), which admits final hyperlong syllables in the output and apparently also retains medial hyperlong configuration -CVVC-C- in cases like (33) in spite of postlexical epenthesis:

\[(34) \text{ ma šif-nā-k-š} ‘we didn’t see you’ (Owens 1984: 158)\]
\[(35) \text{ ma in-fām-š} ‘not swimmable’ (124)\]
\[(36) \text{ dāff ‘has pushed’ (24)\]
\[(37) \text{ gāt-il-ha} \text{ from /gall-l-ha/ ‘he said to her’ (inferred from 33, 105, 113)\]

Perhaps structures like (34) are prosodically accommodated by licensing two semisyllables in a row at the word level.30

The dialect data are very complex, but they seem compatible with the prediction that closed-syllable shortening is bled by lexical epenthesis. Apparent cases of counterbleeding are attributable to independently motivated word-level or postlexical processes that shorten unstressed syllables. This would tend to further support the claim that interaction of phonological processes is transparent within a level and that opacity effects arise from interlevel serialism.

6.7. Summary

I have argued that certain Arabic dialects permit consonants to form unsyllabified moras (onsetless semisyllables) in the word-level phonology. These moras are licensed byadjunction to the prosodic word. Semisyllables persist into the output in certain contexts in C-dialects, in part accounting for the Berber-type syllable structure of these dialects. In VC-dialects an epenthetic vowel is inserted before them as a nucleus postlexically, rendering lexical processes like stress and shortening opaque. Faithfulness requires that the epenthetic vowel be placed before the consonant, so as to maintain its moraic character and to minimize changes in the foot structure of the word. In CV-dialects, a licensing constraint requiring moras to be affiliated with syllables ranks higher, ruling out semisyllables in general. Unsyllabifiable consonants never become moraic but get an epenthetic vowel after them (its place again dictated by prosodic faithfulness). The characteristic syllabic differences between the dialects can largely be explained from these assumptions.

NOTES

1. I present a more extensive treatment of this material in my forthcoming Paradigm Effects and Opacity.
2. But this classification does not coincide with the division between \textit{gil}it and \textit{g}ithu dialects (Jastrow 1978).

3. The words in this table are meant to represent phonological types. Most cited forms are taken directly from the sources cited, but in the Egyptian data I have replaced, with fingers crossed, Behnstedt and Wöllich's type word \textit{silah} 'plowshare' with the corresponding forms of \textit{kitaab} 'book', which are better attested in the other sources.

4. This is a preliminary hypothesis made for the sake of concreteness. The question of where unsyllabifiable moras are adjoined requires more study. On general grounds one would expect them to be adjoined as low as constraints on the form of prosodic constituents allow, since that minimizes violations of Strict Layering. Adjunction to the prosodic word would be motivated by a constraint against the resolution of moraic trochees. A reviewer pointed out that adjunction to a foot entails that unsyllabifiable moras could never occur between two syllables that form a foot (provided adjoinment is allowed only to the edges of constituents). I believe this is true for Arabic in Mohawi and Passamaquoddy, where Hagstrom (1997) argues for adjunction to feet, precisely that structure is motivated. For more on the locus of adjoinment see section 6.2.

5. See Han 1994 for compelling arguments that these are in fact the correct representations and for additional evidence in favor of the To-Mester theory of weak layering.

6. Taking a maximalist position, van der Hulst and Ritter 1998 argue that sesquisyllables consist of two feet, each containing two degenerate syllables.

7. For the moment, let us assume that the foot structure of all dialects of Arabic is organized into moraic trochees.

8. See Behnstedt 1994: 64 for traces of a similar system in the Syrian dialect of Soukhne.

9. The quality of the epenthetic vowel varies but generally in a uniform way for medial, initial, and final syllables in a given dialect. The vowel is most often \textit{i} (in Syrian \textit{a}, sometimes \textit{e} or \textit{i}), and under dialectically varying conditions it is liable to be assimilated to a back vowel in the preceding syllable, e.g., \textit{mnisqutu} 'we'll kill him', \textit{nikub} 'riding', \textit{sahar} 'month' (Palestinian; Palva 1966: 30, 53). Epenthetic vowels are "often pronounced as full vowels and as such are phonetically comparable with phonemic vowels occurring in the same positions" (Palva 1966: 42).

10. Not all VC-dialects undergo "metathesis." Dialects that otherwise have VC-phonology but retain the medial CV-in cases like \textit{yiktabu} are found in Mesopotamia and Anatolia, e.g., Mardin \textit{yektaban} 'they write' (Jastrow 1978: 204), and the Gulf, e.g., Makkan \textit{yiktabu}, \textit{tiktabu} (Jastrow 1980a).

11. The geminate is postexcly shortened in many, perhaps the majority, of dialects; often optionally (Mitchell 1993: 95-96); for some reason no CV-dialect has lexical degemination, which would allow deletion after geminates.

12. Turkish-style final devoicing, without glottalization, is naturally widespread in Anatolia and other Northern dialects of Arabic, including those of the CV-type (Jastrow 1978). Nigerian Arabic glottalizes all pre-pausal segments including vowels (Owens 1983: 21-22), Sudanese Arabic also has final glottalization, perhaps a substructure effect (Reedford, 1996 et al).

13. Nevertheless, there are some VC-dialects in which epenthesis is lexical under some conditions, particularly in the case of clusters that violate sonority sequencing. In

14. A reviewer suggests that the lexicon must form a s

15. Interestingly, Lin's (1 Lin assumes that which seems incom

16. Final -CR clusters are not categorically

17. Similarly the Egypti

18. It appears that there

19. According to Palva (in my terms - P. K.) and in learned baner

20. Epenthetic vowels are parts of Upper Egypt

21. For example, in eastib, unlike other woi

22. Cohen (1965) considers

23. For example, in recot

24. Long -\textit{aa}- is shortened

25. In the Mesopotamia to Jastrow (1978: 212)

26. Interestingly, this cla

27. All those endings trig \textit{h}ai \textit{reb}, however

28. Farwanc (1993) says suffixes like -\textit{r} are unextend to the other v

29. Allophonically and/or v
SYLLABLES AND MORAS IN ARABIC

177

these dialects, a constraint requiring sonority peaks to be syllable heads has been promoted in the lexical phonology.

14. A reviewer suggests that a parallel to this constraint can be found at the level of syllables. There are languages that permit single unfooted syllables but force two adjacent unfooted syllables to form a foot. Similarly, in Arabic, two adjacent unyllabified moras must form a syllable, forcing openness of a nucleus.

15. Interestingly, Lin's (1977) analysis depends on cyclic constraint evaluation. However, Lin assumes that all segments are redundantly moraic in underlying representations, which seems incompatible with Richness of the Base.

16. Final -CR clusters of rising sonority are avoided under some conditions, even if they are not categorically excluded (Cohen 1975: 80).

17. Similarly the Egyptian dialects of the eastern Delta region (Behnstedt and Woidich 1985, map 3).

18. It appears that there is no covert contrast in those contexts in which openness is obligatory, i.e., basically in clusters of rising sonority.

19. According to Palva (1966: 16), “the pronunciation of a consonant as syllabic [moraic, in my terms – P.K.] is often heard in [Lower Galilean Palestinian] in elevated style and in learned borrowings. In colloquial, however, an epenthetic vowel is pronounced before a syllabic consonant.”

20. Epenthetic vowels count for stress in Soqkhun Syrian (Behnstedt 1994) and in some parts of Upper Egypt (Behnstedt and Woidich 1985, map 3).

21. For example, in eastern Libyan they are systematically stressed on the initial syllable, unlike other words: habal ‘footh’ but kitaab ‘he wrote’.

22. Cohen (1963) considers the transitional vowels to be phenomena of phonetic implementation, endorsing Cantineau's statement that “phonologiquement, elles n'ont aucune existence réelle et sont tenues pour zéro par les sujets parlants qui souvent n'ont pas conscience de les prononcer” (Cantineau 1946: 179). Nevertheless, Cohen notes that the transitional vowel can be phonetically identical to phonemic a, though it never gets stressed in environments where a regular vowel would be stressed.

23. For example, in recordings by the Algerian singer Dahmane el Harrachi.

24. Long -ae- is shortened to i or u, or, in dialects that neutralize high vowels, to a. Historically, these long vowels are derived from -VCV- roots by loss of a medial glide or laryngeal, whose color originally determined the quality of the shortened root vowel.

25. In the Mesopotamian qolli dialects it is an optional allophonic process, according to Justrow (1978: 212-213).

26. Interestingly, this class of Egyptian CV-dialects (Behnstedt and Woidich 1985, map 74) includes all those dialects that lengthen all vowels under stress, apparently nonphonemically (Behnstedt and Woidich 1985, map 5): /bagara/ bagara ‘cow’, gafar ‘cat’.

27. All these endings trigger shortening in all CV-dialects, of course; e.g., Egyptian /raah-l-hu/ raahla ‘he went to her’.

28. Farwaneh (1995) accounts for the data in (33) by supposing that sonorant consonantal suffixes like -a are underlyingly moraic, an ad hoc solution that in any case does not extend to the other variations on the shortening theme.

29. Allophonic and/or variable shortening of unstressed vowels is attested for other dialects: see, e.g., Fleisch 1947-1948: 60 on Zaafar (Lebanon), Johnson (1979) on
Palestinian, and Cohen (1975: 55) on “half-long” vowels in Moroccan Jewish Arabic.

30. Syrian shows variation: /zaab-l-na/ zaablauna ‘he brought to us’ (Cowell 1964: 481), /ma-k'-too-b-lak/ makablak ‘written to you (m.)’ (Grotzfeld 1963: 42) or /ma kaan-S nakan/ (Grotzfeld 1980: 178), Soukhne Syrian maas bi-suufis ‘he doesn’t see anything’ (Behnstedt 1984: 162). Iraqi shows stress variation in ease of ephenthesis, e.g. /? bbn-ha – bbn-ha ‘her son’ (Ercin 1963: 41); no shortening is reported for forms like /dabla/, /dabilta – dabilta/ (Ercin 1963: 143). I was unable to find consonantal word-level suffixes like -S in Iraqi to test the prediction.

REFERENCES


SYLLABLES AND MORAS IN ARABIC


SYLLABLES AND MORAS IN ARABIC

7

Semisyllables

Young-mee

7.1. Introduction

Syllables have played a brief period following of syllables has incr. prosodic phonology: a: Smolensky 1993, McF (1990), there can be Universality, which re Exhaustivity, which re or less received view exhaustive is sometin with extreme colloca instance, analyzes Bel as not licensed by syl

(1) Apparent violati
c'tak's'c' 'he ar
q'pisstiz' 'taste

On the other extreme as shown in (2). Hym: that the properties of than the syllable, such not be exhaustive, eve

Earlier versions of this shi Conference at the Universi
during the LSA Summer Science Foundation Works at the Ohio State Universi are particularly indebted to Przepiorki; Martzen 1