George, Leland (1980). Analogical generalization in natural language syntax. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.
Huang, C.-T. James (1982). Logical relations in Chinese and the theory of grammar. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.
Kayne, Richard (1984). *Connected and Binary Branching*. Dordrecht: Foris.
Kuno, Susumu (1980). A further note on Tonoike's intra-subjectivization hypothesis. In MIT Working Papers in Linguistics 2, Yukio Otsu and Ann Farmer (eds.), 171–184. Cambridge: Department of Linguistics and Philosophy, Massachusetts Institute of Technology.
Lasnik, Howard and Mamoru Saito (1984). On the nature of proper government. *Linguistic Inquiry* 14: 235–289.
—— (1992) *Move-α*. Cambridge, MA: MIT Press.
Law, Paul (1991). Verb movement, expletive replacement, and head goverment. *The Linguistic Review* 8: 253–289.
May, Robert (1979). Must COMP-to-COMP movement be stipulated? *Linguistic Inquiry* 10: 719–725.
Rizzi, Luigi (1991). *Relativized Minimality*. Cambridge, MA: MIT Press.
Saito, Mamoru (1985). Some asymmetries in Japanese and their theoretical implications. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.
Saito, Mamory and Keiko Murasugi (1993). Subject predication within IP and DP. Unpublished manuscript, University of Connecticutt and Kinjo Gakuin University.
Takahashi, Daiko (1993). Barriers, relativized minimality, and economy of derivation. Paper presented at the symposium 'On the Minimalist Program', held during the 65th General Meeting of the English Literary Society of Japan, Tokyo.
Takano, Yuji (1992). Proper binding: a minimalist approach. Unpublished manuscript, University of California, Irvine.

Licensing of empty nuclei: The case of Palestinian vowel harmony[1]

SHOHEI YOSHIDA

1. Introduction

This paper attempts to provide a principled explanation of vowel harmony and three closely related phenomena, metathesis, vowel syncope and epenthesis, in Palestinian Arabic, from the perspective of Government Phonology (Kaye, Lowenstamm and Vergnaud 1985, 1990; Charette 1990, 1991). It will be shown that these phenomena can be given a unified account by means of "principles and parameters". Such a principle–parameter approach contrasts markedly with the previous "rule-based approach" on vowel harmony in Palestinian Arabic by Kenstowicz (1981) and Abu-Salim (1987), where vowel harmony, metathesis, syncope and epenthesis are viewed as separate phenomena and expressed by means of re-write rules which are extrinsically ordered.

This paper is organized as follows. Section 2 presents a syllabic analysis of Arabic that does not recognize heavy syllables. This analysis is crucial to my account of vowel harmony, metathesis, syncope and epenthesis in Palestinian Arabic. Section 3 provides an examination of necessary data, while sketching out the previous treatment of vowel harmony. A government-based analysis of Palestinian Arabic is given in section 4. Section 5 discusses a prospect for future studies.

1. An earlier version of this paper was presented at the London Phonology Seminar, January 1992. I am grateful to Monik Charette, John Harris, Yong Heo, Bruce Ingham and Jonathan Kaye, and an anonymous reviewer for *The Linguistic Review*, who gave me valuable comments and suggestions. I am indebted to Muhamad Guerssel and Jean Lowenstamm, who first inspired me to the CVCV analysis of Arabic. Thanks are also due to Sam Kent, Duck-Yong Lee, Kitty Travis and Yuko Yoshida, who contributed to the make-up of this paper. Any errors or inadequacy of analysis are my own responsibility.

The Linguistic Review 10 (1993), 127–159

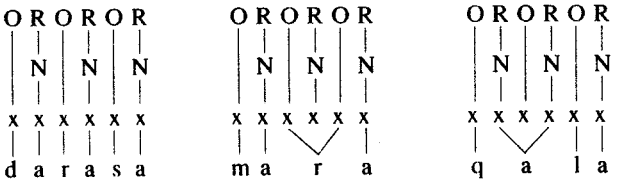
0167–6318/93/0010–0127

2. The syllable structure of Arabic

Following the syllabic analyses of Classical Arabic first proposed by Guerssel and Lowenstamm (1988, 1990) and Lowenstamm (1988), I adopt the assumption that many varieties of Arabic, if not all, have only open light syllables, in contrast to the standard generative analysis (e.g. McCarthy 1979), which recognizes heavy open syllables (i.e. long vowels and diphthongs) and closed syllables. In the "CVCV" analysis of Arabic, which is couched in Government Phonology, long vowels and diphthongs are represented as segments linked to two successive nuclei. Likewise, consonant clusters are represented not as coda–onset clusters but as sequences of successive onsets which are separated by an intervening empty nucleus, as discussed extensively in Kaye (1990a, 1990c). Further, Government Phonology claims that a "word-final" consonant in any language is not a coda closing the word-final syllable, but an onset followed by an empty nucleus, as discussed in section 4.4.[2]

Given this framework, the Classical Arabic verbs *darasa* 'he studied', *marra* 'he passed' and *qaala* 'he said', for example, are analyzed as sharing an identical canonical form, CV.CV.CV, as in (1).[3]

(1)



As one advantage of the CVCV analysis, we need one and only one syllable template for all types of Binyan I verbs, regardless of whether they are triliteral (*darasa*), biliteral (*marra*), or 'hollow' (*qaala* < *qawala*). Compare this with the standard analysis, which has to postulate three different templates, CV.CV.CV, CVC.CV and CVV.CV. Note also that the verb forms cited above are non-pausal forms. Prepausally, final vowels are either lost completely or reduced to a weak schwa, e.g. *darasa* → *daras*, *marra* → *marr*, *qaala* → *qaal*, *kataba* → *katabə* 'he wrote', *radda* → *raddə* 'he replied', *saaqa* → *saaqə* 'he drove'.[4] For those pausal verb forms whose final vowels are deleted, the standard analysis is forced to propose resyllabification, CV.CV.CV → CV.CVC, CVC.CV → CVCC and CVV.CV → CVVC, with the possible extrametricality of the final consonants. Government Phonology adopts a phonological version of the Projection Principle (Kaye, Lowenstamm and Vergnaud 1990), according to which syllable structures (and governing relations holding in these structures) defined at the level of lexical representation remain intact throughout a derivation: a segment that is lexically specified as occupying an onset cannot be reassociated with a coda at the surface level. Resyllabification such as the one above is not expressible and not permitted in Government Phonology. Instead, *darasa* and *daras*, for instance, are claimed to have the same trisyllabic structure, the only difference being whether or not the final nuclear position has segmental content (*da.ra.sa* versus *da.ra.sØ*).

2. See Kaye (1990a) for arguments against the occurrence of codas in word-final position.
3. I leave open the possibility of long-distance (transconsonantal) spreading of the vowel melody /a/, as it is not directly relevant to the present discussion.
4. The reduction rather than deletion of the pausal vowel after certain consonants is a prescribed rule in the Quranic recitation, the most prestigious source of the pronunciation of Classical Arabic. A weak central vowel is heard immediately following the release of the five consonants /q/, /ṭ/, /b/, /ḍ/, /d/ either in pause or followed by another consonant.

Once a language is analyzed as having no consonant clusters, where a cluster is defined as some segment(s) occupying strictly adjacent phonological positions, vowel syncopation and epenthesis must be reanalyzed accordingly. In the standard generative phonology, vowel syncopation is viewed as a process which deletes a vocalic segment together with the position it is associated with, which is followed by resyllabification of the first of the two surrounding consonants to the coda, e.g. (Palestinian Arabic) *fi.him* 'he understood' → *fih.mu* (< *fihim+u*) 'they understood', or to the onset, e.g. (Palestinian Arabic) *ši.rib* 'he drank' → *šrib.na* (< *širib+na*) 'we drank'. In contrast, in the CVCV analysis, Arabic is thought to have open light syllables alone, so this definition of vowel syncopation makes no sense, since the two members of the consonant cluster created as a result of the syncopation of the intervening vowel cannot be viewed as adjacent. Nor does vowel epenthesis make sense in the CVCV analysis. In the standard theory, vowel epenthesis is defined as a process which inserts a vocalic segment along with a skeletal position to break up impermissible consonant clusters, e.g. (Palestinian Arabic) *da.ris* (< dars) 'study'. Such a process does not have theoretical status in the present syllabic analysis of Arabic, since there are no consonant clusters, as defined above, to break up to begin with.

In Government Phonology vowel syncopation and epenthesis are viewed as involving neither the deletion nor creation of nuclear positions during derivations, but are viewed as two different phonetic interpretations of empty nuclei, as demonstrated explicitly in Nikiema (1990), Charette (1991) and Kaye (1990a, 1990b). Given these theoretical perspectives, it should be clear now that new light may be shed on syncope and epenthesis in Palestinian Arabic.

3. Background

In this section I shall present the problems in Palestinian Arabic, while simultaneously introducing the previous rule-based studies. Let us first look at (2). This paradigm of imperfective verbs illustrates vowel harmony and metathesis.

(2)

	subjunctive		indicative	
	a.	b.	a′.	b′.
1 sg.	*ʔá-drus*	*ʔa-rúuħ*	*bá-drus*	*ba-rúuħ*
1 pl.	*nú-drus*	*n-rúuħ*	*b-nú-drus*	*bi-n-rúuħ*
2 sg.m.	*tú-drus*	*t-rúuħ*	*b-tú-drus*	*bi-t-rúuħ*
2 sg.f.	*tú-durs-i*	*t-rúuħ-i*	*b-tú-durs-i*	*bi-t-rúuħ-i*
2 pl.	*tú-durs-u*	*t-rúuħ-u*	*b-tú-durs-u*	*bi-t-rúuħ-u*
3 sg.m.	*yú-drus*	*y-rúuħ*	*bú-drus*	*bi-rúuħ*
3 sg.f.	*tú-drus*	*t-rúuħ*	*b-tú-drus*	*bi-t-rúuħ*
3 pl.	*yú-durs-u*	*y-rúuħ-u*	*bú-durs-u*	*bi-rúuħ-u*
	'study'	'go'		
	c.	d.	c′.	d′.
1 sg.	*ʔá-fham*	*ʔá-lbis*	*bá-fham*	*bá-lbis*
1 pl.	*ní-fham*	*ní-lbis*	*b-ní-fham*	*b-ní-lbis*
2 sg.m.	*tí-fham*	*tí-lbis*	*b-tí-fham*	*b-tí-lbis*
2 sg.f.	*tí-fham-i*	*tí-libs-i*	*b-tí-fham-i*	*b-tí-libs-i*
2 pl.	*tí-fham-u*	*tí-libs-u*	*b-tí-fham-u*	*b-tí-libs-u*
3 sg.m.	*yí-fham*	*yí-lbis*	*bí-fham*	*bí-lbis*
3 sg.f.	*tí-fham*	*tí-lbis*	*b-tí-fham*	*b-tí-lbis*
3 pl.	*yí-fham-u*	*yí-libs-u*	*bí-fham-u*	*bí-libs-u*
	'understand'	'wear'		

3.1. *Vowel harmony*

Let us first inspect vowel harmony. Both Kenstowicz (1981) and Abu-Salim (1987) view vowel harmony as a process by which high front vowels acquire the autosegmental feature [+round] in the vicinity of high rounded vowels. The subjunctive forms in (2a, b, c, d) consist of obligatory prefixes and roots, which may be followed by the 2nd person singular feminine suffix *-i* or the 2nd person and 3rd person plural suffix *-u*. The 1st person singular prefix is invariably *ʔa-*, while the other prefixes are variable. In (2b) they appear as *n*, *t* and *y*, without a following vowel. In (2a), (2c) and (2d) *n*, *t* and *y* are followed by a vowel: they are followed by *u* in (2a) and by *i* in (2c) and (2d). Closer examination reveals that the prefixal vowel is the high rounded vowel *u* if the root contains the same vowel, otherwise, the prefixal vowel is *i*. This indicates that the root vowel *u* is the determinant of regressive vowel harmony.

Vowel harmony is not limited to verbs. The following nouns illustrate progressive vowel harmony.

(3) a. *ʔákil* 'food' *ʤísir* 'bridge' *šámiʕ* 'wax'
b. *fúrun* 'oven' *ʕúrus* 'wedding' *šúɣul* 'job'

(3a) shows that when the initial vowel is either *a* or *i* the second vowel is *i*, while (3b) shows that when the initial vowel is *u* the second vowel is also *u*. It should be mentioned that the underlined high vowels in these nouns are epenthetic, which we will turn to shortly. It is now possible to make the initial generalization that the high front vowel *i* turns into *u* in the environment of *u*.

Next, consider the indicative forms listed in (2a′, b′, c′, d′). Putting aside the precise underlying representation of the indicative prefix for a moment, it is obvious that it begins with the labial stop *b*. In the 3rd person singular masculine and the 3rd person plural indicative forms in (2a′), *bú-drus* and *bú-durs-u*, the prefixal consonant is followed by *u*. On the other hand, in their equivalents in (2c′) and (2d′), *bí-fham*, *bí-fham-u*, *bí-lbis* and *bí-libs-u*, the initial bilabial consonant is followed by *i*. So it appears that the same vowel harmony holds in these forms, too. However, there seem to be some exceptions to our generalization. Let us look at the indicative forms in (2b′). Although the root *-rúuħ* does contain the rounded vowel, as in the root of the verb 'to study' *-drus* in (2a′), the prefixal vowel does not undergo vowel harmony. In fact, the rounded vowel in the so-called hollowed verbs of the type -CuuC and the biliteral root verbs of the type -CuCC never triggers the vowel harmony of the preceding vowel, as seen in (4).

(4) *bi-tšúuf* **bu-tšúuf* *bi-tzúur* **bu-tzúur*
'you (SG MASC) are seeing' 'you (SG MASC) are visiting'
bi-tmúss **bu-tmúss* *bi-trúdd* **bu-trúdd*
'you (SG MASC) are kissing' 'you (SG MASC) are responding'

3.2. *Metathesis*

We will now look at the metathesis found in the triliteral verbs whose root vowel is either *i* or *u* in (2). Their root shape alternate between CCVC and CVCC. More specifically, the unsuffixed forms all have the root shape CCVC, whereas the suffixed forms have the shape CVCC unless the root vowel is *a*. Some alternating forms are extracted below.

(5)

unsuffixed (CCVC)	suffixed (CVCC)	
tú-drus	*tú-durs-i*	*yú-durs-u*
tí-lbis	*tí-libs-u*	*yí-libs-u*

The same metathesis also operates in nouns, as follows:

(6) *fílfil* 'pepper' *fílift-i* 'my pepper'
míšmiš 'apricots' *míšimš-u* 'his apricots'

To summarize, the root shape CCVC is metathesized to CVCC when a vowel-initial suffix is attached and when the root vowel is high: when the root vowel is the low vowel *a* the metathesis is not triggered, as exemplified by the following forms extracted from (2c).

(7) *tí-fham-i (*tí-fahm-i)* *tí-fham-u (*tí-fahm-u)*
*yí-fham-u (*yí-fahm-u)*

Another possible interpretation of this process is found in Kenstowicz (1981), where the metathesis process is broken into two processes, syncope and epenthesis. In his analysis the vowel in CCVC is first syncopated, yielding a triconsonantal cluster CCC, and then the vowel reappears between the initial and second consonants, yielding CVCC. An advantage of this analysis is that one need not recognize the existence of a metathesis process in Palestinian Arabic in addition to syncope and epenthesis, which exist independently of metathesis.

3.3. *Syncope*

The syncope process in Palestinian Arabic deletes certain vowels in nonfinal open syllables. As shown in (8), vowels which are deleted are high vowels.

(8)				
	nízil	'he descended'	*nízl-u*	'they descended'
	ʔákil	'food'	*ʔákl-u*	'his food'
	šúɣul	'job'	*šúɣl-i*	'my job'
	fúrun	'oven'	*fúrn-i*	'my oven'

However, stressed high vowels escape syncope, as illustrated by (9).

(9) *nzílit *nizlit (nizilit ← nizil + t)* 'I descended'

The underlying form of the verb is *nizil + t*. When the epenthesis vowel *i* appears to break up the final root consonant *l* and the following suffix *t*, the form has two open syllables, *ni* and *zi*, both of which meet the structural description of syncope. However, the initial vowel is deleted and the second vowel remains, rather than the other way round, due to the fact that the second vowel is assigned stress prior to syncope. Alternatively, one might claim that syncope applies prior to stress assignment, deleting the initial vowel, and then the remaining second vowel is stressed. However, such an analysis is implausible because all perfective 1st person singular forms of the triliteral verbs have stress on the second syllable, irrespective of whether or not the initial vowel is deleted, e.g. *darásit* 'I studied', *katábit* 'I wrote'. Note that long vowels *ii* and *uu* are not deleted even though they are high.

The syncope rule in Kenstowicz (1981) is reproduced in (10).

(10) Syncope rule

$$\begin{bmatrix} +\text{syll} \\ +\text{high} \\ -\text{long} \\ -\text{stress} \end{bmatrix} \rightarrow \emptyset / \_\_\_\_ \text{ CV}$$

However, Kenstowicz notes that the structural description of this rule requires some additional provisos.

First of all, the low vowel *a* may be deleted if it is in non-initial position and the vowel of the following syllable is also *a*, as in (11).

(11) -CaCaC (3 SG MASC) verbs lose the second *a* when followed by *-at* (SG FEM)

3 SG	3 SG FEM	3 PL	
dáras	*dárs-at*	*dáras-u *dárs-u*	'studied'
ħtáram	*ħtárm-at*	*ħtáram-u *ħtárm-u*	'respected'
n-kásar	*n-kásr-at*	*n-kásar-u *n-kásr-u*	'broke'

The final *a* of *dáras*, *ħtáram* and *n-kásar* is deleted if and only if the following suffixal vowel is also an *a* as in the 3rd person singular feminine forms; it is not deleted when followed by the plural suffix *u* as in the 3rd person plural forms.

Second, and interestingly, for a vowel to be deleted, the immediate preceding syllable must be light, as pointed out by Kenstowicz (1981).

(12)			
	sáabak-at	'raced'	*(*sáabk-at)*
	dárras-at	'studied'	*(*dárrs-at)*
	tikáttir	'you (SG MASC) increase'	
	tikáttir-u	'you (PL) increase'	*(*tikattr-u)*

In *sáabak-at* the stem-final *a* is followed by the identical vowel *a* in the suffix. But the stem-final vowel unexpectedly escapes syncope because the preceding syllable contains a long vowel. In *dárras-at* the stem-final vowel is not deleted because the preceding syllable is closed by the first member of the geminate.

Finally, the initial high front vowel in the indicative verbal forms does not undergo syncope when the following root shape is either -CVVC or -CVC1C2 (C1=C2), as in *bi-rúuħ* 'he is going' and *bi-rúdd* 'he is responding'. It should be remembered that not only does the initial high front vowel in these indicative forms escape syncope, it is also immune from vowel harmony as mentioned above.

3.4. *Epenthesis*

In the great majority of Palestinian Arabic words, biconsonantal clusters are permitted medially, but no clusters are permitted prepausally. When impermissible clusters, that is, medial triconsonantal clusters or prepausal biconsonantal clusters, are created in the underlying representation, the high front vowel *i* is inserted in the positions given in (13).

(13) Ø → i / C__CC or C__C#.
šuf-t il-maktuub 'I saw the letter'
šu šufi-t 'what did you (SG MASC) see?'
šufi-t maktuub 'I saw a letter'

The examples here all contain the verb 'I saw', which is created by suffixing *t* to *šuf* (< *šaaf*). In *šuf-t il-maktuub* the *f-t* cluster is in medial position, so it surfaces phonetically as such. In *šu šufi-t* the *f-t* cluster is not permitted as it is in final position, and so the epenthetic vowel appears between them. In *šufi-t maktuub*, the epenthetic vowel appears to prevent the creation of the triconsonantal cluster *ftm*.

3.5. *The rule-based treatment*

Having observed all the processes, we are ready to examine the previous works more closely. In the previous works by Kenstowicz (1981) and Abu-Salim (1987), vowel harmony is viewed as a process by which the feature [+round] is mapped on to all high vowels in the stem, which is defined as the vowel harmony domain. In the case of verbs, the stem consists of a prefix and a root. All suffixes are excluded from the vowel harmony domain as suffixal vowels do not undergo vowel harmony, and nor do they cause the preceding vowels to undergo vowel harmony. Kenstowicz's (1981) treatment is illustrated in the following derivation of *tu-ktub-i* 'you (SG FEM) write', extracted from Abu-Salim (1987).

(14)
[ti-ktib] i	underlying representation
+r	
[tí-ktib] i	stress assignment
+r	
[tí-ktb] i	syncope
+r	
[tí-kitb] i	epenthesis
+r	
[tí-kitb] i	vowel harmony = mapping
⋁	
+r	
[túkutbi]	

The prefixal high vowel and epenthetic vowels are treated as unrounded and high (i.e. *i*) in their underlying representations, since they are realized as *i* when they occur in the context of *a* or *i*, or outside the vowel harmony domain. The high vowel in the roots is treated as being unspecified for the feature [round] and represented as *I*. All the high vowels within the harmonic domain are realized as *u* when the root is associated with the feature [+round] as in (14), otherwise they appear as *i*.

Couched in distinctive feature theory, the previous analysis above requires two types of partially specified representations for the high front vowel in the prefixes and the high front vowel in verb roots, though their phonetic realizations are identical: *i* in the prefixes would be represented as [+hi, –round], whereas *i* in verb roots would be represented as [+hi]. It should be noted that the prefixal and epenthetic vowels, which are specified as unrounded in their underlying representations, are converted into rounded vowels if the roots are associated with [+round], requiring the replacement of [-round] with [+round]. The undesirability of the rule-based analysis above hardly needs pointing out. Rules, which take the formalization A -> B / C__D, are purely descriptive and do not explain why syncope, epenthesis and vowel harmony take place as such. It is also noteworthy that the rule-based treatment in (14) necessarily requires arbitrary rule ordering: vowel harmony syncope feeds epenthesis and epenthesis is ordered before vowel harmony to ensure the prefixal vowel undergoes vowel harmony.[5]

4. A government-based approach

4.1. *The Empty Category Principle*

While I am in complete agreement with the previous studies on the point that stress placement takes place prior to the other processes, I argue against the necessity of treating vowel harmony, metathesis, syncope and epenthesis as separate phenomena. In what follows I will show that a government-based analysis enables us to unify these phenomena into a single process, eliminating the need for writing three different rules and arbitrarily ordering them in a satisfactory fashion. In fact the principle–parameter approach requires no rules, hence no rule ordering.

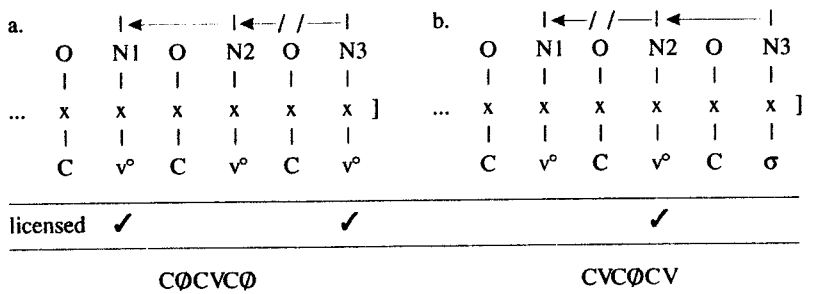
Government Phonology claims that the phonetic interpretation of empty nuclei is controlled by the following principles:

5. Abu-Salim fuses the syncope rule and the epenthesis rule into a single metathesis rule at the cost of having to recognize metathesis, syncope and epenthesis as three different processes.

(15) a. The Empty Category Principle (ECP) (Kaye 1990a)
   i) A licensed empty nucleus has no phonetic realization.
   ii) An empty nucleus is licensed if it is properly governed or if it is domain final in languages which license domain-final empty nuclei.
   b. A nuclear position a properly governs a nuclear position β iff
   i) α is adjacent to β on its projection.
   ii) α is not itself licensed.
   iii) No governing domain separates α from β.

What these principles state is that for a nonfinal empty nucleus to receive no phonetic interpretation it must be licensed by an adjacent nucleus with segmental content which properly governs it. The patterns of vowel/zero-alternation in many languages show that proper government is head-final. In other words, an empty nucleus is inaudible when there is an audible nucleus to its right. When a nonfinal empty nucleus fails to be licensed it is phonetically manifested. In contrast, final empty nuclei do not require an external licenser to receive zero phonetic realization; individual languages choose one of the available parametric values (*yes* or *no*) as to whether or not final empty nuclei are licensed. For example, Arabic, Chinese and English license final empty nuclei, and accordingly words in these languages may end with a consonant followed by an inaudible empty nucleus, while Italian, Japanese and Swahili do not license final empty nuclei and therefore words in these languages must end with a phonetically realized nucleus. Consider (16).[6]

(16) (←— = proper government)



Two hypothetical configurations are provided where all nuclei are empty, except the final nucleus of structure (16b), which has a lexical segment indicated by σ.

6. Empty nuclear positions are filled by the cold vowel v°. See Kaye, Lowenstamm and Vergnaud (1985) for details.

In (16a), the final empty nucleus N3 is licensed by virtue of the fact that this language licenses final empty nuclei, so it need not be pronounced. Being licensed itself, N3 cannot properly govern N2, so N2 must be phonetically realized. Being unlicensed, N2 can properly govern N1, so N1 need not be phonetically expressed. This yields the surface pattern CCVC, which in fact contains two empty positions (CØ.CV.CØ). In (16b) the final nucleus N3 does have a segment, so it can license N2. Being licensed, N2 is not pronounced, and cannot license N1, so N1 is pronounced, yielding the surface form CVCCV, which is in fact CV.CØ.CV.[7]

Let us see how the ECP is implemented in the real world. I propose that there are three ways by which the cold vowel (*v*°) in unlicensed empty nuclei obtains phonetic realization.

First, there are languages where the cold vowel is pronounceable. Kaye, Lowenstamm and Vergnaud (1985) propose in their element-based segmental theory that empty positions that are not filled by real elements are in fact filled by the cold vowel *v*°. Thus the cold vowel is synonymous with the absence of any of the three main vocalic elements *I*, *U* and *A* proposed in Government Phonology.[8] As the elements *I*, *U* and *A* represent "palatality", "roundedness" and "pharyngeality", respectively, or, in terms of the distinctive features, [−back], [+round] and [−high], the phonetic value of the cold vowel is [+back, −round, +high]. The cold vowel pronounceable in isolation is found in Moroccan Arabic (Kaye 1990b) and Korean (Lee and Yoshida 1992) for example. In these languages, unlicensed empty nuclei surface as unrounded high back vowels, as in (17).

(17) Moroccan (*u*/Ø):
lØbusØ 'he wore' *lubØs-u* 'they wore'
Korean[9] (*u*/Ø):
mitu-mØ 'believe+nominalizer' *mitØ-k'o* 'believe+and'

Secondly, there are languages in which the cold vowel is not pronounceable in isolation. These languages provide a conflicting situation where unlicensed empty nuclei must be pronounced, as required by the ECP, but the cold vowels

7. This is not to deny the existence of true closed syllables. Government Phonology defines that a true coda, which is a governed position, is obligatorily followed by an onset that governs it. The consequence of the coda–onset transsyllabic government is that a sequence of a true coda and a following onset always exhibit severe phonotactics.
8. The elements *I*, *U*, *A* in isolation are phonetically realized as [i], [ʊ], and [ɑ] respectively. Acoustic evidence of elements are discussed in Williams and Brockhaus (1992). Elements are normally associated with superscripts that indicate the charm values of the elements (I°, U°, A^{+}). I omit the superscripts in this paper as they are not essential to the following discussion.
9. In fact, an additional principle is needed for empty nuclei to be licensed in Korean as argued in Lee and Yoshida (1992).

occupying these empty nuclei cannot be pronounced. The unpronounceable cold vowels can be rendered pronounceable only by acquiring some elements in one of the two following ways. In the first case, added elements have no discernible sources, and are dubbed "ambient elements" by Kaye, Lowenstamm and Vergnaud (1990). Different languages choose different ambient elements, as illustrated in (18).[10]

(18) Hindi (*ə*/Ø):

pičØk-a	'squeezed'	*pičəkØ*	'squeeze'

Yawelmani (i/Ø):

paʔØt-alØ	'fight (dubitative)'	*paʔitØ-ka*	'fight (imperative)'

Tangale (*u*/Ø):

tarØgo	'trap'	*tarugØno*	'my trap'

Typically, a single element is added to the cold vowel in these cases.[11]

In the second case, the unpronounceable cold vowel in unlicensed empty nuclei obtains phonetic realization by acquiring some element(s) existing within its domain.[12]

(19) Slovak (*e*/Ø, *o*/Ø):

pasØm-o	(GEN PL)	*pasemØ*	'fiber (NOM SG)'
bahØr-a	(GEN SG)	*bahorØ*	'wheelrim (NOM SG)'

As proposed in Kaye (1990c), "yers" in Slavic languages are also subject to the ECP. The examples in (19) display alternations *e*/Ø and *o*/Ø. I suggest that *e* and *o* are part of the underlying representations of these words, and that they are detached from licensed nuclei and attached to unlicensed nuclei to satisfy the ECP. Another example of the phonetic realization of the cold vowel by means of obtaining an available element comes from Japanese. As mentioned earlier, final empty nuclei are not licensed in Japanese, so all Japanese words are expected to end with a phonetically realized nucleus. However, not all Japanese words end with a vowel: words may also end with the so-called moraic nasal, e.g. *hoN* 'book', *seN* 'line'. What is the syllabic status of such a nasal? It cannot be a coda because words in all languages end with a nucleus, as per the Coda Licensing Principle. At the same time, we cannot say that the parametric value of final empty nuclei selected by Japanese (i.e. the unlicensed nature of final empty nuclei) is relaxed when the preceding onsets contain a nasal, for a parametric value must be upheld throughout a system. The problem disappears

10. The Tangale examples are taken from Nikiema (1990).
11. Each of the epenthetic vowels *ə*, *i* and *u* in (18) is composed of a single element, A+, I° and U°, respectively. For a detailed discussion of the internal structure of each sound, see Kaye, Lowenstamm and Vergnaud (1985).
12. Slovak examples are taken from Kenstowicz and Kisserberth (1979).

if we assume that the unpronounceable cold vowel in a lexically empty final nucleus obtains a nasal segment that spreads from the preceding onset, as proposed in Yoshida (1990).[13]

(20)

[hoũ]

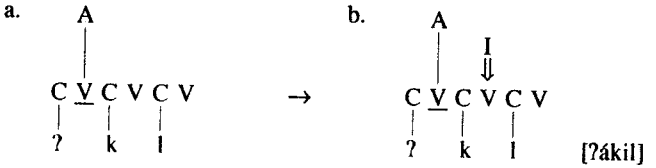
The cold vowel obtains phonetic realization by combining with the nasal element, yielding a nasalized cold vowel. Recall that the theoretically predicted phonetic value of the cold vowel (or the absence of *I*, *U*, and *A*) is unrounded high back. This prediction is borne out, as the nasalized cold vowel appears as a nasalized unrounded high back vowel, *ɯ̃*.

4.2. *Epenthesis and syncope*

I claim that Palestinian Arabic makes use of both ambient elements and available elements. I propose that only unstressed nuclei can be properly governed in Palestinian Arabic. The cold vowel in unlicensed nuclei is phonetically expressed by obtaining the element *U* if there is one in its domain; otherwise it obtains the ambient element *I*. Thus, among the three elements *I*, *U* and *A* which appear in nuclear positions in Palestinian Arabic, only *U* has the property of spreading.

The derivation of the noun *ʔákil* in (3a) is given below.

(21) (V=stressed)



[ʔákil]

13. Compare this with Itô's (1986) analysis. Assuming that word-final consonants in languages are codas and that the word-final nasal in Japanese is a consonant, Itô proposes a coda filter that permits only a nasal and excludes other consonants in a word-final coda. Such a solution is purely descriptive and the fact (as interpreted in Itô) that why a nasal may occupy a word-final coda but other consonants cannot is treated as being accidental.

Skeletal points usually adopted in Government Phonology are replaced by Vs and Cs here in order to place vowels and consonants on different tiers for expository convenience. In (21a) only the initial nucleus is morphologically associated with the element *A* and the other two nuclei are empty. As Arabic licenses final empty nuclei, the final nucleus can remain empty. However, the medial nucleus is not licensed by the licensed final nucleus, so it must obtain phonetic content in one way or another. There is one element in this domain, *A*, but as it cannot spread in Palestinian Arabic, the medial nucleus has recourse to an ambient vowel, *i*, yielding *ʔákil*.

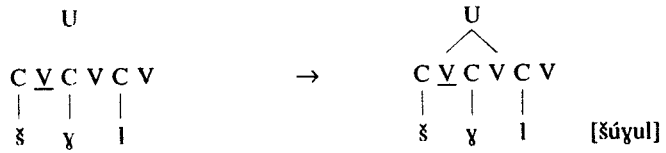
Next we shall consider the derivation of *ʤísir* in (3a).

(22)

The underlying structure above has no nuclear segment at all. Therefore, unlicensed nuclei can only be phonetically expressed with recourse to ambient elements. Starting from the right-hand side of the structure, the final nucleus is licensed in domain-final position, therefore the medial nucleus is not licensed. The initial nucleus cannot be licensed as it is stressed. Consequently, ambient vowels are added to the unlicensed initial and medial nuclei.

(23) shows the derivation of *šúɣul* in (3a).

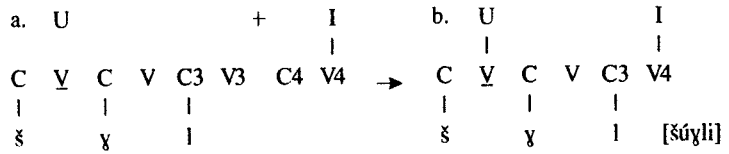
(23)



For the same reasons as in (22), the unlicensed initial and medial nuclei must obtain some phonetic content. In this structure, there is a floating vowel, *u*, which may dock onto a nucleus or nuclei to satisfy the ECP. Thus the initial and medial nuclei are associated with this element, yielding *šúɣul*.

When a vowel-initial suffix is attached to these nouns, the final high vowel of each noun is deleted. This is due to the fact that the second nucleus of these nouns is licensed by the final nucleus that obtains a suffixal vowel. (24) is the derivation of *šúɣli* 'my job' (< *šúɣul* + *i*).

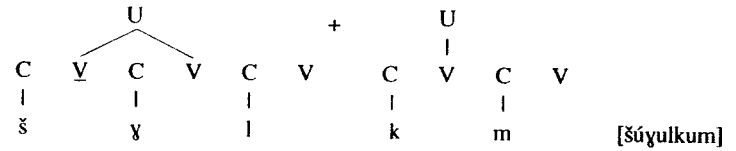
(24)



The suffix is composed of a syllable of which nucleus is morphologically linked to the *I* element. This suffix is attached to a noun whose final empty nucleus is licensed, creating an underlying sequence of a vowelless syllable C3V3 followed by an onsetless syllable C4V4. It is claimed in Government Phonology that the Obligatory Contour Principle (OCP: Leben 1973; McCarthy 1986 and Yip 1988, etc.) operates on two adjacent nuclei under certain circumstances, in addition to its more general operation on melodic tiers. In a sequence of a vowelless syllable and an onsetless syllable above, that is $[\sigma]_{O1}[\emptyset]_{N1}[\emptyset]_{O2}[\sigma]_{N2}$, the OCP always operates on the first empty nucleus and the following nucleus with a segment that are not separated by an intervening onset consonant. As a result the two syllables are fused, yielding $[\sigma]_{O1}[\sigma]_{N2}$. This particular application of the OCP will be discussed in depth in a later section. In the resulting structure (24b), the penultimate nucleus no longer needs to obtain phonetic realization, since this position is licensed by the final nucleus which now has a segment, producing *šúɣli*.

When a suffix begins with a consonant, the final empty nucleus of unsuffixed noun-stems remains empty, and accordingly, the nucleus preceding the empty nucleus must be phonetically manifested, as shown in the representation of *šúɣulkum* 'your (pl.) job' (< *šúɣul+kum*).

(25)



The suffix-final empty nucleus is licensed by virtue of the fact that Arabic licenses domain-final empty nuclei. The antepenultimate nucleus is licensed by the penultimate nucleus containing a segment, and therefore cannot license the second nucleus. The unlicensed second nucleus and the stressd initial nucleus are phonetically manifested by obtaining *U*.

4.3. *Vowel harmony*

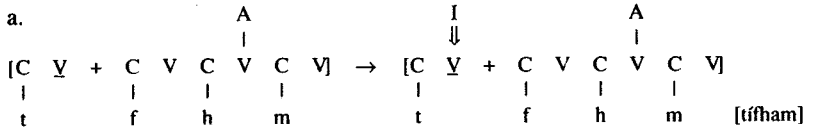
Equipped with the theoretical preliminaries, let us proceed to analyze vowel harmony in the verbal forms in (2). I assume that the vowel harmony domain consists of the imperfective prefixes, which appear in the subjunctive forms in (2a, b, c, d), and the following roots. Only the 1st person singular prefix *ʔa* has a vowel. The other prefixes have the following representations where the onset is occupied by a segment, followed by an empty nucleus.

(26)

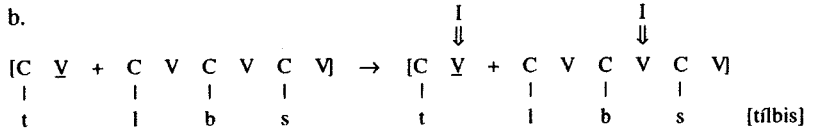
Consider the derivations of *tífham*, *tílbis*, and *túdrus*.

(27)

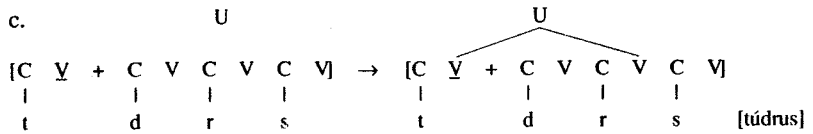
a.



b.



c.

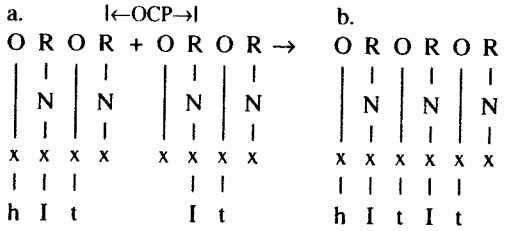


These structures consist of four nuclei of which the final nucleus is empty. As domain-final empty nuclei are licensed in Arabic, the licensing pattern for the four nuclei, starting from the rightmost nucleus, is licensed-unlicensed-licensed-unlicensed, leaving the initial and the penultimate nuclei unlicensed. Being unlicensed they must be phonetically realized in one way or another. In the underlying representation of *tífham* in (27a) the penultimate nucleus is morphologically linked to an element, *A*, so the ECP is already satisfied (the reason why *A* is not floating will be given later). The unlicensed initial nucleus must acquire an ambient element, *I*, since the *A* element does not have the property of spreading. The verb *tílbis* has no underlying vowels at all, and so both the initial and penultimate nuclei acquire ambient elements. The underlying form in (27c) differs from the previous two cases in that it has the element *U* which may spread and therefore may dock onto any nucleus when required by the ECP. It is linked to the initial and penultimate nuclei, generating the form *túdrus*.

4.4. *Syllable superimposition*

Before analyzing metathesis, we shall examine a related process which I call *syllable superimposition*, since this process will be an integral part of the following analysis. Recall that we considered a special case of the OCP, where a vowelless syllable and a following onsetless syllable are obligatorily collapsed to one syllable, as exemplified by the phrase *hit it*.

(28)

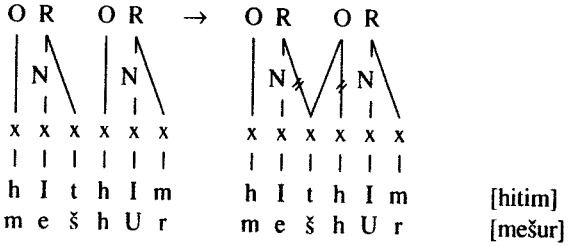


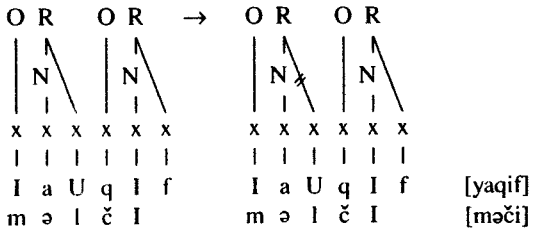
The third syllable in (28a) is superimposed on the second syllable in a metaphorical sense. In (28a), the second nucleus is empty. However, superimposition may also operate on two nuclei both of which contain segments if the two nuclei are not separated by a consonant, as in the case of the syncopation of a vowel in the French definite article when followed by a vowel-initial noun, for example *le+ami* → *lami* 'the friend' (see Charette 1991). What interests us are the following forms.

(29)
a. *hit him* → *hitim*
b. *mešur* → *mešur* 'celebriyu' (Turkish)
c. *yawqif* → *yaqif* 'he stops' (Classical Arabic)
d. *məl+či* → *məči* 'far; connective' (Korean)

The standard interpretation of these processes is to assume resyllabification. (29a) and (29b) exhibit the deletion of the initial segment of the second syllable, which is followed by resyllabification. In (29c) and (29d), the final segment of the initial syllable is deleted. Thus the standard analysis of (29) would be as follows:

(30) Standard analysis





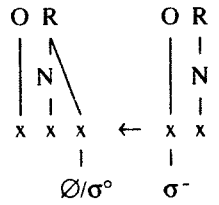
The initial syllables of the underlying forms are all treated as closed syllables. This is an expected state of affairs in the standard generative phonology where the first segment of a (phonetically observed) word-internal biconsonantal cluster is always thought to occupy coda position, unless the cluster forms a branching onset. Note that in the standard analysis, (29a, b), (29c, d), and the syncopation of one of two consecutive vowels, such as in the case of the French definite article, are viewed as three unrelated phenomena. In Government Phonology these processes may be unified under the notion of superimposition. Consider the following principle:

(31) Coda Licensing Principle (Kaye 1990a: 311)

A post-nuclear rimal position must be licensed by a following onset.

This means that whenever a rime branches, the "post-nuclear rimal point", or what is traditionally called coda, must be licensed by the following onset which governs it, as shown below.[14]

(32)



The superscripts indicate the charm values of segments. Details apart, the governed coda must be either empty (the case of gemination) or filled with a "charmless" segment (σ°), a segment that involves neither stiff vocal cords nor slack vocal cords, i.e. nasals, glides and "charmless" obstruents (obstruents other than fully voiced obstruents and fortis obstruents). On the other hand, the governing onset point must be filled either by a (negatively) charmed segment

14. The term *coda* has no theoretical status in Government Phonology as coda is not considered to be a syllable constituent, as argued by Kaye, Lowenstamm and Vergnaud (1990).

(σ⁻) which involves either stiff or slack vocal cords, that is fortis obstruents and fully voiced obstruents, or by a charmless segment which is equal to or more complex than its governee.[15] A significant implication of this principle is that domain-final consonants never occupy post-nuclear rimal position, as there is no following onset that licenses them: rimes may branch only word-internally. Even word-internally, the occurrence of a post-nuclear rimal position is possible if and only if it is linked to a potential governee (e.g. liquids, nasals and glides, etc.) and the following onset is linked to a potential governor (e.g. fortis stops and fully voiced stops, etc.). Thus, in Government Phonology, severe phonotactic constraints on the two members of coda–onset clusters are claimed to be a universal property of languages.

Applying the Coda Licensing Principle to the words in (29), *hit* cannot be analyzed as consisting of a single closed syllable because the final *t* can in no way be licensed, as there is no governing onset to its right. Instead, Government Phonology claims that *hit* has the representation *hitØ* where *t* occupies an onset which is followed by an empty nucleus. Similarly, the initial syllable of *mešhur* cannot be analyzed as being closed, either, since *š* and *h* cannot constitute a well-formed trans-syllabic cluster: *h* is not a potential governor but a governee, and *š* is not a potential governee but a governor.[16] The sequence of *š* followed by *h* never forms a coda–onset cluster in languages. In fact the two consonants are not adjacent; they are separated by an empty nucleus, i.e. *mešØhurØ*. The Arabic imperfective form *yaqif* derives from the underlying form *yawqif*. The existence of *w* in the underlying representation is supported by the perfective form *waqaf* 'he stopped'. Compare this with the imperfective and perfective forms of the normal triliteral verb 'to write' *yaktub* and *katab*. Imperfective forms of normal triliteral verbs are formed by two processes: prefixing (*ya* = 3rd person masculine) and converting the perfective canonical shape C1VC2VC3 into C1C2VC3. If two segments σ1 σ2 can occupy C1 and C2 of imperfective forms respectively, then their mirror image σ2 σ1 may also occupy C1 and C2 respectively. Recall that if C1 and C2 formed a coda–onset trans-syllabic cluster then segments that may occupy these positions would have to show asymmetry. However, the complete lack of distributional constraints on C1 and C2 in imperfective verb forms proves the absence of a governing relation between C1 and C2, i.e. C1 and C2 cannot be analyzed as a coda and an onset. The only plausible analysis of C1 and C2 is that they are separated by an empty nucleus. Thus the underlying form *yawqif* has the syllable structure *ya.wØ.qi.fØ*. Finally, sequences of two non-identical consonants in Korean are also separated by an

15. I follow Harris (1990) who proposes the Complexity Condition, which states: "Let α and β be segments occupying the positions A and B respectively. Then, if A governs B, β must be no more complex than α."

16. For detailed discussions of the governing properties of segments, see Harris (1990), Kaye, Lowenstamm and Vergnaud (1990).

empty nucleus for independent reasons that do not conern us here.[17] Thus the underlying form *məlci* has the representation *məlØci*.

To sum up, the processes in (29) are not different from the French vowel syncope in *le+ami* → *lami*. The processes in (29) differ from the French case only in that one of the two nuclei involved in superimposition is empty, as shown in (33).

(33) Syllable superimposition

a.

O1	N1	O2	N2	O3	N3	O	N	
x	x	x	x	x	x	x	x	
h	I	t		h	I	m		[hitim]
m	e	š		h	U	r		[mešur]
I	a	U		q	I	f		[yaqif]
m	ə	l		č	I			[məči]

b.

O1	N1	O3	N3	O	N	
x	x	x	x	x	x	
h	I	t	I	m		[hitim]
m	e	š	U	r		[mešur]
I	a	q	I	f		[yaqif]
m	ə	č	I			[məči]

(33a) shows that O3 and N3 are superimposed on O2 and N2. In the case of *hit it* , O2 has *t* but O3 is empty, and so the superimposition of O3 upon O2 does not result in a segmental clash. A problem arises when both O2 and O3 have segments. (33a) shows that either O2 or O3 has a charmless segment, or very informally speaking, a "weak" consonant, and it is this segment that deletes when it would otherwise clash with another consonant. Note that syllable superimposition does not involve any resyllabification as defined above, unlike the standard treatment where syllable structures may be distorted freely. In the present analysis, segments that are lexically linked to onsets end up being linked to onsets throughout a derivation.

4.5. *Metathesis*

Having seen syllable superimposition, let us proceed to an analysis of metathesis in verb forms of Palestinian Arabic. The metathesis of the root pattern (CCVC → CVCC) takes place only before vowel-initial suffixes. In addition, the root vowel *a* escapes metathesis. Compare the following forms from (2): *yi-lbis* versus *yi-libs-u*, *yu-drus* versus *yu-durs-u* but *yi-fham* versus *yif-ham-u* **yi-fahm-u*. Let us consider the derivation of *yi-libs-u*.

17. See Lee and Yoshida (1992) for a detailed analysis of consonantal clusters in Korean.

(34)

a. [C V̲ + C V C V C V]C V → C V̲ + C V C V C V

I I b s I I b s [yílibsu]

b. I ⇓ I ⇓ U

In the underlying form (34a) all the nuclei in the stem are empty. Notice that the stem-final syllable has an onset filled with *s* followed by an empty nucleus, whereas the plural suffix has an empty onset followed by a nucleus filled with *U*. Thus nothing prevents the suffixal syllable from being superimposed on the stem-final syllable. The resultant structure (34b) has three empty nuclei and the final nucleus is filled with *U*. The penultimate nucleus is licensed and therefore inaudible by virtue of the fact that it is properly governed by the final nucleus, which has now obtained *U*. The second nucleus is unlicensed and must therefore be phonetically realized. The first nucleus is stressed, so it cannot be licensed, either. As a result, the initial and the second nuclei are phonetically manifested with recourse to the ambient element *I*, yielding the form [yílibsu].[18]

Next, let us consider *yu-durs-u*.

(35) U U U U

[C V̲ + C V C V C V]C V → C V̲ + C V C V C V

I d r s I d r s [yúdursu]

As in the previous example, the first and second unlicensed nuclei must obtain some means of phonetic realization. Unlike the previous example, however, the floating and spreadable element *U* is available within the harmonic domain. This element is linked to the two unlicensed nuclei to satisfy the ECP, yielding [yúdursu].

Finally, let us consider the form *yi-fham-u* which resist metathesis.

(36)

a. [C V̲ + C V C V C V]C V → C V̲ + C V C V C V

I f h m I f h m [yífhamu]

A U b. I ⇓ A U

18. We need to understand why the suffixal vowel cannot spread to the unlicensed nuclei in the stem, despite the fact that it can license the stem-final nucleus. I have no ready answer to this question.

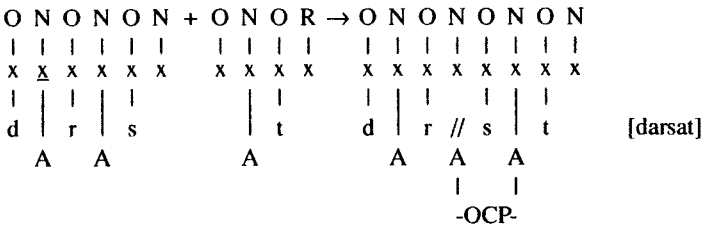
As a result of syllable superimposition, the final nucleus in (36b) has a segment and is therefore a potential governor. We would expect that it properly governs the preceding nucleus and that *A* delinks from its skeletal point, as was seen in the Tangale example in (18), *tarØgo+no* → *tarugØno*. However, this does not take place: the *A* element is not deleted. As the penultimate nucleus is unlicensed, it can license the second nucleus. Being stressed, the initial nucleus is unlicensed anyway. Consequently, the initial nucleus alone obtains the ambient element, generating *yífhamu*. Recall that I proposed earlier that *A* is morphologically linked to verb roots. This claim is supported by the fact that if *A* was floating, then we would end up with the incorrect form **yífahmu*.

4.6. *The OCP and government-licensing*

I have attributed the absence of metathesis in *yífhamu* above to the fact that the root vowel *a* cannot be deleted upon suffixation. However, the root vowel *a* is unexpectedly deleted if the suffixal vowel is also *a*. I assume that this is an effect of the OCP complementing the ECP.[19]

Let us consider the underlying representation of *dars-at* 'she studied' which is derived from *daras* + *at*.[20]

(37)



19. The deletion of a governee when it is identical or homorganic to its governor is a common phenomenon. When this applies to either a post-nuclear rimal point (coda) or the first of two successive onsets, the result is gemination; when it applies to the governed position within a branching nucleus or to the first of two successive nuclei, the result is a long vowel.

20. I am aware that the standard generative analysis of Semitic melodies is adherent to the assumption that all identical vowels and consonants within roots, e.g. (Arabic) takaatab 'he corresponded', (Arabic): sammam 'he poisoned', are manifestations of a single segment associated with two or more positions. I am not convinced by this unprincipled and excessive employment of the OCP whereby the surface identical sounds ..σ..σ..σ.. within roots are always underlyingly related with each other. I assume that, except in cases of geminates and long vowels, there are σ's corresponding to the surface ..σ..σ..σ.. at the level of lexical representation, rather than a single σ, and that the applicability of the OCP to given σs is subject to universal principles such as government-licensing (Charette 1990) and the ECP.

The OCP clears away the second *A* to provide an environment for proper government. Being licensed, the second nucleus need not be phonetically realized.

Recall that the deletion of the stem-final *A* is blocked if it is preceded by a "heavy" syllable. Two examples are extracted from (12).

(38) *dárras-at* **dárrs-at* *sáabak-at* **sáabk-at*

The question to ask is what blocks the deletion of the stem-final *a*. Consider first *dárras-at*. The answer to this problem comes from the concept of *government-licensing* proposed in Charette (1990: 241). Let us deviate from Palestinian Arabic for a moment and examine government-licensing closely.

(39) Government-licensing

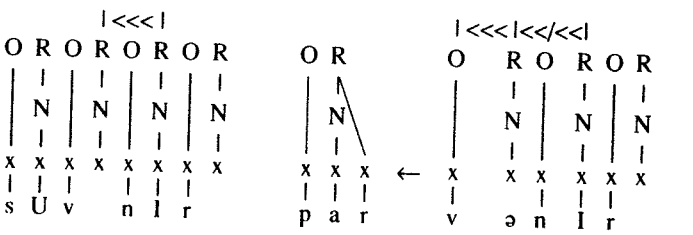
For a governing relation to hold between a non-nuclear head A and its complement B, A must be licensed to govern by its nucleus at the licenser projection level. The government-licenser of an onset is an unlicensed nucleus.

The point of this principle is most easily appreciated by comparing the following two French words taken from Charette (1991).

(40) *souvenir* [suvØnirØ] *parvenir* [parvənirØ] *[parvØnirØ]

Charette proposes that empty nuclei in French are inaudible when licensed but are realized as schwas when unlicensed. In *souvenir* the second nucleus, which is empty, is licensed by the following nucleus containing *i*, and therefore is inaudible. In *parvenir* we have the same sequence *vØni*, but the empty nucleus is expressed as a schwa. According to Charette, what distinguishes the two forms is the fact that in *parvenir* the empty nucleus is preceded by a non-nuclear point, containing *v*, which governs the rimal complement of the initial branching rime, containing *r*, as per the Coda Licensing Principle. Charette claims that as the governor *v* must be government-licensed by the second nucleus, a schwa is manifested in this position, although the second nucleus is followed by an unlicensed nucleus which would otherwise properly govern it, as illustrated below.

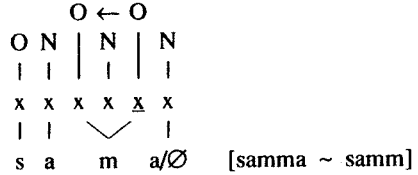
(41)



In some languages domain-internal empty nuclei and domain-final empty nuclei behave differently with respect to government-licensing. Again take French as

an example. Although internal empty nuclei cannot government-license the preceding onset head to govern the coda of the preceding syllable from right to left, or the complement in a branching onset from left to right, final empty nuclei have the ability to government-license these governing onset heads, e.g. [kaRtØ] 'card' (where the final empty nucleus licenses the preceding onset to govern the coda), [katRØ] 'four' (where the final empty nucleus licenses the preceding onset head to govern its complement within a branching onset). In fact, empty nuclei in many Arabic dialects, including Palestinian Arabic, have these properties, too, as seen in Guerssel and Lowenstamm's (1988) analysis of Classical Arabic metathesis.[21] They argue that in triliteral verbs such as *samm-a* (non-pausal) ~ *samm* (pausal) 'he extended' the third onset governs the second one.

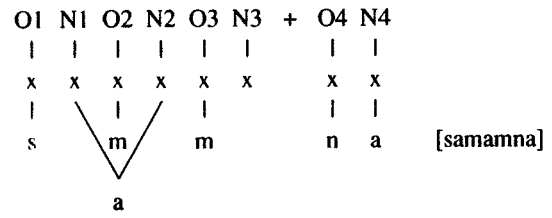
(42)



[samma ~ samm]

I assume that the so-called biliteral roots are in fact triliteral roots whose second and third consonants are identical (e.g. √*smm*), and that these identical consonants are fused into one by the OCP if and only if they can constitute a governing domain. In other words, in a geminated consonant the second position must be government-licensed by the following nucleus. In (42) the third onset is government-licensed by the final nucleus, irrespective of whether or not it is empty. Therefore, the OCP applies to the underlying two *m*s, forming a geminate. Following Kaye (1990a), I assume that an empty nucleus contained within an inter-onset governing domain is licensed by the domain, not by a nucleus outside the domain. The suffixation of consonant-beginning morphemes induces metathesis. For example, attaching the 1st person plural morpheme *na* to the stem *samm* does not result in **samm-na* but *samam-na*, as in (43).

(43)



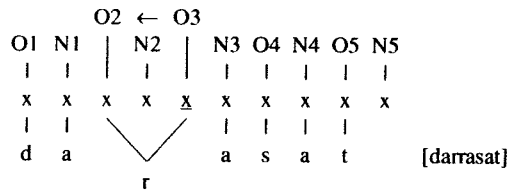
[samamna]

21. In some Arabic dialects, geminates are degeminated prepausally. In our account, this measns that final empty nuclei in these dialects are not government-licensers of the preceding governing onsets.

As a result of the suffixation, the stem-final empty nucleus N3 is now domain-internal. Unlike in domain-final position, it can no longer government-license O3. Not being government-licensed, O3 cannot govern O2 to form gemination, hence the two identical consonants *m*s remain as such. Consequently, N2 is no longer licensed: it is not contained within an interonset governing domain, and nor can it be licensed by N3 which is licensed by N4. Therefore N2 must be phonetically spelled out. It obtains *a* spreading from N1.

Returning home, the reason why proper government fails to apply to the second *a* in *darrasat* is parallel to Guerssel and Lowenstamm's analysis of metathesis above.

(44)



[darrasat]

As N3 and N4 both contain the same segment *a*, in normal circumstances N4 would properly govern N3 as an OCP effect. However, the gemination of *r* requires the government of O2 by O3. If N3 was properly governed (licensed) then it could not government-license O3 to govern O2. Thus the proper government of N3 by N4 is blocked to preserve gemination, rendering the form **darrsat* ill-formed. The observant reader might wonder why government-licensing has priority over proper government in this form, unlike the case of *samamna* in (42) where proper government had priority over government-licensing. In fact these facts are not arbitrary. Unlike the root √*smm* above, the Binyan II verb *darras* has three different root consonants √*drs*, and so even if N3 is properly governed by N4, the form like **dararsat* can never be produced.

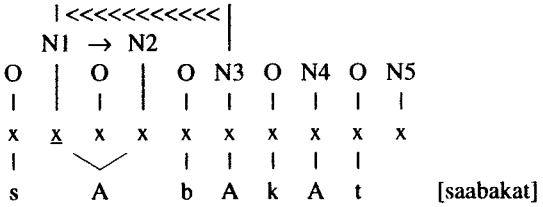
The stem-final nucleus is also prevented from being properly governed if the preceding vowel is long, e.g. *saabak-at *saabk-at*. We can account for this case by extending the concept of government-licensing to the head of long vowels. It is proposed in Yoshida (1992) that in Choctaw, Luganda and Palestinian Arabic, for a nuclear head to govern its complement, the head must be government-licensed by the following nucleus. Recall that in French the government-licenser of a governing onset head is either an unlicensed nucleus or a licensed final empty nucleus. This is also true with the government-licenser of a governing nuclear head in these languages, as exemplified by the Palestinian Arabic examples in (45).

(45) a. *beetØ* 'a house' *beet-i* 'my house'
ruuħØ 'go! (SG MASC)' *ruuħ-u* 'go! (PL)'
b. *betØ-na* 'our house' *ruħØ-li* 'go for me!'

(45a) shows that the head of each long vowel can be government-licensed by the following final nucleus, irrespective of whether or not the final nucleus has a segment. In (45b) long vowels are systematically shortened before an internal empty nucleus, due to the fact that internal empty nuclei are not government-licensers. It is interesting to point out that there are no word-final long vowels in Palestinian Arabic (*VV##).[22] This is an expected state of affairs, since there is no government-licenser to the right of a word-final vowel. Kaye (1990a) has shown that long vowels are systematically shortened before empty nuclei in Turkish and Yawelmani. The same vowel shortening is found in Choctaw and Luganda as well. Interestingly, one thing common to all these languages is the fact that they have "pseudo" consonant clusters, whose members cannot be analyzed as adjacent at the skeletal level, as per the Coda Licensing Principle, and, in the case of Luganda, on tonal facts. Finally, it is extremely interesting to point out that all these languages, like Palestinian Arabic, do not allow final long vowels, a rather surprising coincidence.

Let us consider the underlying representation of *saabakat*.

(46)



The long vowel *aa* is represented as a sequence of two nuclei N1 and N2 which form a head-initial internuclear governing domain. It must be head-initial as the second element of an apparent long vowel never bears stress. The head of this nuclear governing domain must be licensed by the following unlicensed nucleus N3. Therefore, N3 cannot be licensed by N4, and consequently, is phonetically manifested. One might wonder why the verb form above chooses the retention of the long vowel rather than the proper government of N3 by N4 and the consequent vowel shortening. I assume that the vowel is not shortened in this Binyan III verb so that it does not take the same form as its Binyan I counterpart *sabØk-at* 'she preceded'.

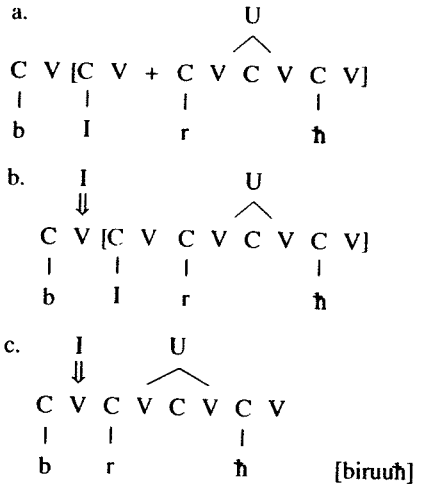
4.7. *The indicative forms*

Finally, we will examine the indicative forms in (2). The anomaly of the indicative forms is that the prefixal vowel harmonizes with the root vowel *u* only in normal triliteral verbs and not in hollow verbs, e.g. *bú-drus*, *bú-durs-u* versus

22. Apparent word-final long vowels are followed by *h*, e.g. *fihimtuu(h)* 'they understood him'.

bi-rúuħ, *bi-rúuħ-u*. It should also be recalled that the prefixal vowels also resist vowel harmony when the root is the type CVC1C2 (C1=C2) as in *bi-rúdd*. I propose that the indicative prefix is outside the harmonic domain and that the *u* that appears after *b* at the surface level is in fact a part of the person/gender prefix. (47) is the derivation of *bi-rúuħ*.

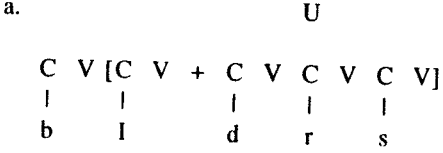
(47)



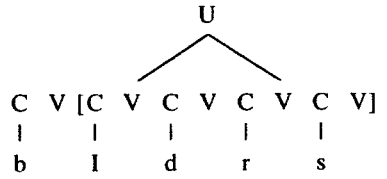
As seen in (47a), the indicative prefix is composed of an onset filled with *b* and an empty nucleus. The root vowel *u* is linked to the initial and second nuclei of the root. Having segmental content, the root-initial nucleus can properly govern the empty nucleus of the person/gender prefix. The nucleus of the indicative prefix cannot be licensed since the following nucleus is licensed. Thus the nucleus of the indicative prefix must be phonetically manifested. As it is outside the harmonic domain, it cannot obtain the root vowel *u*, and as such is the case, it obtains the ambient vowel *i*, as in (47b). This underlying form *biyØruuħ* is reduced to *biruuħ* by the superimposition of the root-initial syllable on the person/gender prefix *yØ*.

Next, consider the derivation of *búdrus*, which shows vowel harmony.

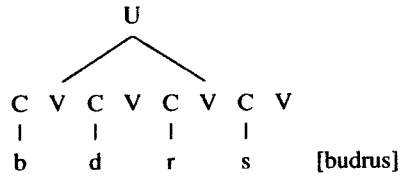
(48)



b.



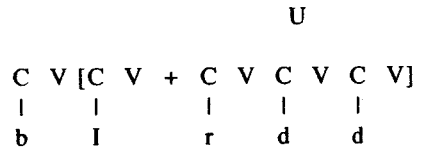
c.



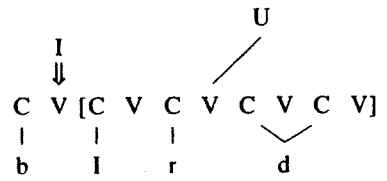
Unlike in (47), the *U* element is not lexically associated with any nucleus in the underlying form (48a). Therefore, its association is subject only to the ECP as in (48b). (48c) is the surface form where the second syllable is superimposed on the initial syllable.

Finally, (49) is the derivation of *bi-rúdd*.

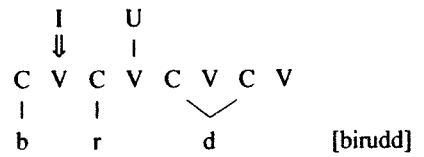
(49) a.



b.



c.



The underlying representation (49a) shows that the root contains two identical consonants and a floating *U*. As Palestinian Arabic is a language where final empty nuclei can government-license preceding onsets, the two *d*s are collapsed into one, forming a head-final interonset governing domain. The empty nucleus contained within the interonset governing domain is licensed by the domain.

Being licensed, the penultimate cannot properly govern the preceding nucleus, which is root-initial. The root-initial nucleus cannot be properly governed by the final nucleus, either, as a governor and its governee must not be separated by an intervening governing domain. Therefore, the root-initial nucleus is phonetically spelled out by obtaining the floating *U*. The nucleus of the person/gender prefix need not be pronounced as it is properly governed by the following nucleus which now has *U*. On the other hand, the nucleus of the indicative prefix is not licensed, and accordingly needs to be phonetically realized. As it is outside the harmonic domain it acquires the ambient element I as in (49b). Finally, the third syllable in (46b) is superimposed on the second one, and thus the surface structure (49c) is derived.

5. A prospect for future studies

I have shown that vowel harmony, metathesis, syncope and epenthesis in Palestinian Arabic are manifestations of the ECP. While a rule-based approach only describes phonological processes, principle-based approaches, such as this one, searches for an explanation. The phonological phenomena in Palestinian Arabic are not treated as language-specific, but as manifestations of the principles and parameters also attested in other languages. Note also that the present analysis requires no arbitrary resyllabification. That is, while other frameworks are concerned with creating and deleting syllable constituents, Government Phonology is concerned with principled phonetic interpretation of syllable constituents. What clearly distinguishes the present analysis of Arabic from the standard analysis is the CVCV syllabic analysis, which has not yet been accepted by most linguists. It should be pointed out, however, that the validity of the standard syllabic analysis of Arabic has almost never been questioned by its proponents. Let us examine the nature of the standard syllabic analysis in some depth.

All previous generative studies on Arabic syllable structure have been built on the following assumptions, which are also applied to other languages. Each assumption is exemplified by Classical Arabic words.

(50) a. C in the word-final string -VC is a coda of the final syllable, e.g. *qul* 'Say! (SG MASC)', *ʔuk.tub* 'Write! (SG MASC)'

b. Long vowels in languages are invariably represented as a segment linked to two "tautosyllabic" prosodic units, e.g. *laa* 'no', *baab* 'door (pausal)'.

c. C1 in the string -VC1C2V is the coda of the first syllable, unless C1 and C2 constitute a branching onset of the second syllable, e.g. *yak.tu.bu* 'he writes', *rad.da* 'he replied'.

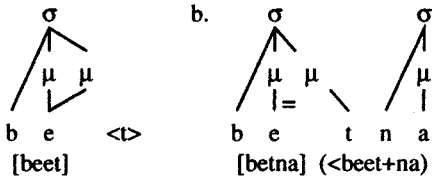
To my knowledge, no attempt has been made to verify these assumptions. While evidence in support of (50a) is absent, evidence against (50a) abounds. Harris (1992) has pointed out some properties of word-final consonants, which are not shared by genuine word-internal codas. He stated:

> A word-final consonant behaves like an internal onset, not like an internal coda.
> a. In stress assignment, it fails to contribute to the weight of the rime to which it supposedly belongs (extrametricality).
> b. It fails to trigger closed-syllable shortening in the rime to which it supposedly belongs.
> c. Alleged final coda clusters frequently contravene otherwise general sonority sequencing constraints.
> d. The phonotactics of final consonant clusters parallel those of internal coda–onset clusters or branching onsets.

Note that the standard analysis assigns a duel status to a word-final consonant: it is "extrametrical" when it behaves as if it did not belong to the final syllable, otherwise it is a coda. Government Phonology does not need the extrametricality of a word-final consonant, since it is treated as an onset.

Next, let us consider standard assumptions (50b) and (50c). Is it really the case that all long vowels are monosyllabic, and the first member of a bi-consonantal cluster is always a coda (if the two consonants do not belong to an onset)? One might say that these assumptions are supported by the following moraic analysis of vowel shortening in Palestinian Arabic.

(51) a. b.



(51a) shows the word-final consonant *t* is required to be extrametrical. In the suffixed structure (51b), for the initial syllable to retain the bimoraicity of the vowel and contain *t* at the same time, an extra mora is required. As there are only two moras available, the second mora is delinked from the vowel and reassociated with *t*. Although the moraic analysis correctly predicts the shortening of the vowel, it is premised on assumptions (50b) and (50c): it does not prove that they are correct. Given the fact that any obstruent may occupy any position of a word-internal cluster in Arabic, one of the consequences of the assumption in (50c) is the claim that any consonant may occupy any of two adjacent positions by default, reducing phonotactic constraints observed on consonant clusters to language-specific phenomena. In fact, the assumptions in (50) is not based on any analytic work. As there is no reason to take the

standard assumptions in (50) as given, there is no reason why the standard syllabic analysis of Arabic which recognizes monosyllabic long vowels and codas should be preferred to the present CVCV analysis either.

It follows that if the CVCV analysis of Arabic is correct, then all phonological phenomena that have been thought to make references to syllable weight would have be reconsidered. Stress assignment would be one major field in need of reconsideration. All previous generative studies on Arabic stress (e.g. McCarthy 1979; Welden 1980; Halle and Vergnaud 1987; Goldsmith 1990) have relied on a ternary syllable weight distinction: light, heavy and superheavy. As the CVCV analysis does not recognize heavy and superheavy syllables, an account of stress in this framework cannot refer to syllable weight. In what follows, I will briefly sketch a course for future studies on Arabic stress in the CVCV framework, since a full analysis of stress in Palestinian Arabic is of course far beyond the scope of this paper. Consider the two different syllabic analyses in (52).

(52)

	Standard	CVCV	
a.	*ka.táb.ti.*	*ka.tá.bØ.ti.*	'you (SG FEM) wrote'
b.	*mak.túu<b>*	*ma.kØ.tú.u.bØ.*	'letter'
c.	*ka.tá.bi<t>* (< *ka.tab.+t*)	*ka.tá.bi.tØ*	'I/you (SG MASC) wrote'
d.	*ká.ta<b>*	*ká.ta.bØ*	'he wrote'
e.	*ká.ta.bu.*	*ká.ta.bu.*	'they wrote'
f.	*mak.tá.ba.to.*	*ma.kØ.tá.ba.to.*	'his library'

It is important to note that the stress patterns in (52) are far from exhaustive, but are only sufficient to illustrate the difference between the two approaches. The generalization of the limited data above in the standard syllabic analysis would be as follows:

(53)
> a. Stress the rightmost heavy syllable if it is ultimate or penultimate (52a, b, c).
> b. Stress the leftmost syllable not exceeding the antepenult if is not followed by heavy syllable in a phrase (52d, e, f).
> c. Final consonants are marked extrametrical (52b, c, d).
> d. Assign stress to underlying structures before epenthesis (52c).

On the other hand, as far as the data in (52) goes, all we need to say in the CVCV analysis is the following:

(54) Stress the antepenultimate syllable.

What is required is a metrical analysis where metrical constituents are built on empty nuclei as well as nuclei containing segments. This line of research is by no means innovative. Charette (1991) has successfully analyzed French stress by taking empty nuclei into consideration, and Segundo (1990, forthcoming) claims

that stress and related phenomena in Portuguese are best explained with a syllabic analysis that incorporates empty nuclei. I am convinced that the inclusion of empty nuclei in metrical analysis of Arabic, or in other languages for that matter, will open up a new path to fruitful generalization.

Received 12 May 1992
Revised version received
5 June 1993

Japan Centre
Faculty of Asian Studies
Australian National University
Canberra ACT 0200
Australia

References

Abu-Mansour, M. H. (1991). Epenthesis in Makkan Arabic: Unsyllabified consonants vs. degenerate syllables. In *Perspectives on Arabic Linguistics III*, Bernard Comrie and M. Eid (eds.). Amsterdam: John Benjamins.

Abu-Salin, I. M. (1980). Epenthesis and geminate consonants in Palestinian Arabic. *Studies in the Linguistic Sciences* 10: 1–11.

—— (1986). Vowel shortening in Palestinian Arabic. *Lingua 68*: 223–240.

—— (1987). Vowel harmony in Palestinian Arabic: A metrical perspective. *Journal of Linguistics 23*: 1–24.

Charette, Monik (1988). Some constraints on governing relations in phonology. Unpublished Ph.D. dissertation, McGill University, Montreal.

—— (1990). License to govern. *Phonology* 7: 233–253.

—— (1991). *Conditions on Phonological Government*. Cambridge: Cambridge University Press.

Goldsmith, John (1990). *Autosegmental and Metrical Phonology*. London: Basil Blackwell.

Guerssel, Mohammed and Jean Lowenstamm (1988). Classical Arabic 'metathesis'. Paper presented at the Generative Phonology Workshop 'How different is phonology?', University of Leiden.

—— (1990). Stem formation and the derivation of verbal categories in Classical Arabic. Paper presented at the GLOW Colloquium, School of Oriental and Asian Studies, London.

Halle, Morris and Jean-Roger Vergnaud (1987). *An Essay on Stress*. Cambridge, MA: MIT Press.

Harris, John (1990). Segmental complexity and phonological government. *Phonology* 7: 255–300.

—— (1992). The autosegmental effects of prosodic licensing. Paper presented at the Government Phonology Workshop, 7th International Phonology Meeting, Krems.

Itô, Junko (1986). Syllable theory in prosodic phonology. Unpublished Ph.D. dissertation, University of Massachusetts, Amherst.

Kaye, Jonathan (1990a). 'Coda' licensing. *Phonology* 7: 301–330.

—— (1990b). Government in phonology: The case of Moroccan Arabic. *The Linguistic Review* 6: 131–159.

—— (1990c). On the interaction of theories of lexical phonology and theories of phonological phenomena. In *Proceedings of Krems Phonology Conference*, Krems.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud (1985). The internal structure of phonological representations: A theory of charm and government. *Phonology Yearbook* 2: 305–328.

—— (1990). Constituent structure and government in phonology. *Phonology* 7: 193–231.

Kenstowicz, Michael (1981). Vowel harmony in Palestinian Arabic: A suprasegmental analysis. *Linguistics* 19: 449–465.

Kenstowicz, Michael and C. Kisserberth (1979). *Generative Phonology*. New York: Academic Press.

Leben, William (1973). Suprasegmental phonology. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Lee, D. and Shohei Yoshida (1992). Interonset government in Korean. Unpublished manuscript, Australian National University.

Lowenstamm, Jean (1988). 'How similar is syntax?' Paper presented at the Generative Phonology Workshop 'How different is phonology?', University of Leiden.

McCarthy, John (1979). Formal problems in Semitic phonology and morphology. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

—— (1986). Today the world, tomorrow phonology. *Phonology Yearbook* 3: 27–45.

Nikiema, Emmanual (1990). Vocalic epenthesis reanalyzed: The case of Tangale. *Current Approaches to African Linguistics 7*: 41–51. Dordrecht: Foris.

Segundo, Silvia. (1990). On empty nuclei and metrical structure in Brazilian Portuguese. SOAS Working Papers in Linguistics and Phonetics 1990: 39–52, University of London.

—— (Forthcoming). Stress and related phenomena in Brazilian Portuguese. Unpublished Ph.D. dissertation. School of Oriental and Asian Studies, University of London.

Sibawaihi, A.B.A. (1968). Al-Kitab, vol II. Reprint of Bulaq edition. Baghdad: Al-Muthanna.

Welden, Ann (1980). Stress in Cairo Arabic. *Studies in the Linguistic Science* 10: 99–120.

Williams, Geoff and Wiebke Brockhaus (1992). Automatic Speech Recognition: A principle-based approach. In SOAS Working Papers in Linguistics and Phonetics 2: 371–404, University of London.

Wise, Hilary (1981). Some functionally motivated rules in Tunisian phonology. *Journal of Linguistics* 19: 165–181.

Yip, Moira (1988). The obligatory contour principle and phonological rules: A loss of identity. *Linguistic Inquiry* 19: 65–100.

Yoshida, Shohei (1990). A government-based analysis of 'mora' in Japanese. *Phonology* 7: 331–351.

—— (1991). Some aspects of governing relations in Japanese phonology. Unpublished Ph.D. dissertation, School of Oriental and Asian Studies, University of London.

—— (1992). Licensing of nuclear heads. Paper presented at the Government Phonology Workshop, the 7th International Phonology Meeting, Krems.