Levin, Juliette (1985). *A metrical theory of syllabicity*. PhD dissertation, MIT.
Lindsey, Geoff & John Harris (1990). Phonetic interpretation in generative grammar. *UCL Working Papers in Linguistics* **2**. 355–369.
McCarthy, John J. (1988). Feature geometry and dependency: a review. *Phonetica* **45**. 84–108.
Maran, La Raw (1971). Burmese and Jingpho: a study of tonal linguistic processes. *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics* **4**.
Nespor, Marina & Irene Vogel (1986). *Prosodic phonology*. Dordrecht: Foris.
Ní Chasaide, Ailbhe (1989). Sonorization and spirantization: a single phonetic process? Ms, Trinity College, Dublin.
Pike, Kenneth L. (1947). *Phonemics*. Ann Arbor: University of Michigan Press.
Prentice, D. J. (1971). *The Murut languages of Sabah*. Canberra: Australian National University.
Prince, Alan S. (1984). Phonology with tiers. In Aronoff & Oehrle (1984). 234–244.
Rice, Keren D. (1990). On deriving sonority: a structural account of sonority relationships. Ms, University of Toronto.
Sagey, Elizabeth (1986). *The representation of features and relations in non-linear phonology*. PhD dissertation, MIT.
Schaefer, Ronald P. (1981). A strength hierarchy for Tswana. *Studies in African Linguistics* **12**. 114–117.
Schane, Sanford S. (1984). The fundamentals of particle phonology. *Phonology Yearbook* **1**. 129–155.
Selkirk, Elizabeth O. (1982). The syllable. In Harry van der Hulst & Norval Smith (eds.) *The structure of phonological representations*. Part 2. Dordrecht: Foris. 337–384.
Selkirk, Elizabeth O. (1984). On the major class features and syllable theory. In Aronoff & Oerhle (1984). 107–136.
Smith, Norval S. H. (1988). Consonant place features. In Harry van der Hulst & Norval Smith (eds.) *Features, segmental structure and harmony processes*. Part 1. Dordrecht: Foris. 209–235.
Vennemann, Theo (1972). On the theory of syllabic phonology. *Linguistische Berichte* **18**. 1–18.
Vennemann, Theo (1988). *Preference laws for syllable structure and the explanation of sound change*. Berlin: Mouton de Gruyter.
Vennemann, Theo & Peter Ladefoged (1973). Phonetic features and phonological features. *Lingua* **32**. 61–74.

Phonology 7 (1990) 301–33
Printed in Great Britain

*'Coda' licensing**

Jonathan Kaye
SOAS, University of London

0 Introduction

One of the main thrusts of the research programme known as Government Phonology has been the attempt to replace the rule component of a phonology by a group of universal principles common to all linguistic systems along with a series of parameters which delimit the nature of linguistic variation from one system to another. Given the wealth of analyses that employ rules and the fact that their number continues to grow, even to this day, it is a daunting task to find plausible alternative stories for each and every analysis that employs phonological rules in some crucial way. One of the leading ideas of work done on syllable structure that I have participated in for the last ten years has been the emergence of such alternative analyses for a series of phenomena that seemed to call for phonological rules.[1] Slow but steady progress has been made on formulating such principles and showing that many rule-based analyses could be successfully replaced by a syntax-like 'principles and parameters' approach.[2]

One of the earliest attempts to replace a rule-based account by something less arbitrary concerned the commonplace phenomenon of closed syllable shortening, i.e. a long vowel or heavy diphthong would shorten (or not occur) in a closed syllable. This phenomenon presented a challenge for our programme. It was difficult to see at first glance what the relationship was between the process, the shortening of a vowel, and the context in which it occurred, a closed syllable. In this paper I shall return to this phenomenon. I will sketch the history of its treatment and show the influence it has had on the evolution of government phonology. I will draw some conclusions about syllable structure that may, at first glance, seem entirely preposterous. It is my belief that there is an increasing amount of empirical support for the rather strange claims concerning syllable structure that emerge from the study of this phenomenon.

1 So-called 'closed syllable shortening'

Consider the following data from Turkish and Yawelmani:[3]

(1) *Turkish*[4]

NOM	POSS	ABL	NOM PL	
merak	mera:kɨ	meraktan	meraklar	'law'
sevap	seva:bɨ	sevaptan	sevaplar	'good deed'
usul^{y}	usu:l^{y}ü	usul^{y}den	usul^{y}l^{y}er	'method'

Yawelmani

sa:pit	saphin	sapnit	'burn'
go:bit	gobhin	gobnit	'take in'
panat	pana:hin	pana:nit	'arrive'
ʔilet	ʔile:hin	ʔile:nit	'fan'
ʔamlal	ʔa:miltaw	ʔa:milka	'help'
moxlol	mo:xiltaw	mo:xilka	'grow old'

The above examples constitute standard 'workbook' type problems. Each language displays a number of phonological phenomena, but what interests us here is the alternation in vowel length that is found in both languages. In standard generative phonology the 'correct' answer consists of the formulation of a rule of vowel shortening in closed syllables. Note that the initial short vowel in *panat* precludes a putative open syllable lengthening solution. This rule is formulated below.[5]

(2) $V: \rightarrow V / \underset{R}{\_\,C}$

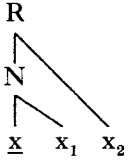
(2) expresses the generalisation that long vowels shorten in closed syllables. As is to be expected in the rule-based approach, there is an arbitrary relationship between the phonological event – the shortening – and the context in which the event takes place – in closed syllables. This 'solution' to the vowel-length alternation problem appeared unsatisfactory for a number of reasons. For one thing, the process is quite widely attested and yet its formulation in a rule-based model seemed no more natural than a host of other processes which were either much rarer or not attested at all. There was no apparent reason why vowels should shorten in such a context. Why shouldn't we find vowels lengthening in closed syllables? The theory gave us no reason to expect a process such as (2) rather than one which shortened vowels in, say, open syllables. Furthermore, at the time Lowenstamm and I were trying to develop an explanatory theory of phonological processes; one in which each phonological event had a cause. The rule-based approach involving processes like (2) seemed particularly inappropriate for such a programme.

Considerations of this sort led Lowenstamm and me to seek a principle from which (2) would follow as a logical consequence. We called this 'the

theory of prosodic government'.[6] In brief, prosodic government suggested that the nucleus, or more exactly, the nuclear position on the skeletal tier, could be viewed as the head of the rhyme. Further, the nuclear head could reasonably be required to govern all other members of this constituent. Government was expressed configurationally in this theory. We required that the nuclear head position C-command all other members of the rhymal constituent. Prosodic government would then have the property of excluding all cases of branching nuclei contained within branching rhymes. In other words, closed syllable shortening would fall out as a logical consequence of prosodic government. To see how this is done consider the structures in (3) below:

(3) a. 

b. 

c. 

In order for A to C-command B, the first branching node that dominates A must also dominate B. Structure (3a) represents a closed syllable. The first branching node that dominates the (underlined) nuclear head is R, the rhyme. This node also dominates the 'coda' position and so prosodic government is satisfied. Structure (3b) represents a long vowel or heavy diphthong. The constituent N immediately dominates the head as well as the governed member of the nucleus. Once again prosodic government is satisfied. Finally, a long vowel or heavy diphthong is shown in (3c). Within the nucleus the head does C-command its sister, x_1. However, in the rhymal constituent the nuclear head does not C-command x_2. The first branching node dominating the head, *viz.* N, does *not* dominate x_2. Structure (3c) is illicit.

Prosodic government then gives us a principled way of accounting for closed syllable shortening. The phenomenon reduces to a consequence of constraints on the possible form of syllables. All this seemed quite encouraging. Since prosodic government was viewed as a part of UG, we were of course required to make that claim that it held for all phonological systems. Thus, closed syllables containing branching nuclei should never be found. On the face of things, this claim seemed to be contradicted by a veritable flood of examples from such familiar languages as English, Arabic, dialects of French, etc. To say the least, it seemed somewhat brash to claim that branching nuclei were universally excluded from closed syllables given forms such as English *keep*, (Quebec) French [ve:r] *vert* 'green', Arabic [ḍa:r] 'house', etc. What was required was to look at a number of these cases and see if the violations of prosodic government that they displayed could be shown to be illusory.

In considering the class of languages which display closed syllable shortening in one form or another, we encountered two types of systems. The first, which we termed 'well-behaved', showed vowel shortening in

exactly the contexts where the theory seemed to predict it. This included both word-internal shortening illustrated by Turkish *meraktan* (< *mera:ktan*) and Yawelmani *saphin* (< *sa:phin*) and final closed syllable shortening as in Turkish *merak* (< *mera:k*) and Yawelmani *ʔilet* (< *ʔile:t*). Thus, Turkish and Yawelmani were two languages that behaved well with respect to closed syllable shortening. Apparently, beyond prosodic government nothing needed to be said about their length alternations.

In contrast to Turkish and Yawelmani an uncomfortably large number of languages appeared to be ill-behaved. As mentioned above, English, (Quebec) French and Arabic were three such languages. They displayed a considerable number of forms where long vowels and heavy diphthongs appeared to occur in closed syllables. These results were somewhat discouraging. Nevertheless, these systems had one property that gave us cause for some optimism in pursuing a principle-based account of closed syllable laxing. Each of these languages manifested some form of this phenomenon. In other words, the presence of a goodly number of forms which seemingly violated prosodic government did not mean that manifestations of this principle were entirely absent in the phonological systems. Thus, alongside *keep*, *child* and *five*, English contains *kept*, *children* and *fifty*. Quebec French has [ve:r] 'green (m.)' but it also has [vert] (*[ve:rt]) 'green (f.)'. So even languages that appeared to violate prosodic government contained some forms that seemed to indicate its presence in their system. In fact, English and Quebec French shared an interesting property with respect to apparent violations of prosodic government; the vast bulk of the apparent exceptions occurred before single word-final consonants.[7] This aberrancy was noted by Charette (1984). More remarkably, she noted that the identical pattern existed in Wolof, a West Atlantic language spoken in Senegal and neighbouring countries. Wolof seemed to permit long vowels in closed syllables but only if the syllable was closed by a single word-final consonant. The following data illustrate this point (Charette 1984: 50f):

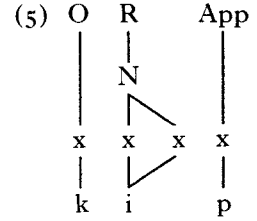
(4)	IMPERFECTIVE	INVERSIVE	
	rɔɔf	rɔppi	'to put in/to take out'
	yɛɛw	yɛwwi	'to tie/to untie'
	tɛɛr	tɛddi	'to start/to stop a vehicle'

In general, the distribution of long vowels follows the pattern of (4). Long vowels are found in open syllables or else in syllables that are closed by a single final consonant. By and large this same pattern is found in Moroccan Arabic.[8]

Let us sum up the situation. There are some languages like Turkish and Yawelmani which manifest vowel shortening in exactly the positions predicted by prosodic government. Long vowels are absent from all closed syllables, including those closed by a single word-final consonant. These are the so-called well-behaved languages. There are a number of languages which seem to manifest long vowels or heavy diphthongs in closed syllables. We can say two things about these languages: (i) the vast bulk

of these exceptional cases involve long vowels occurring before single word-final consonants and (ii) these same languages display the effects of prosodic government in other contexts. Surely the existence of these 'exceptional' cases could not be coincidental. Why should prosodic government be violated so frequently in precisely the environment of a final syllable closed by a single consonant? Charette attempted to come to grips with this question by using a notion introduced by Halle & Vergnaud (1978) to deal with an entirely different problem: the skewing of consonant distribution in Germanic languages in word-final position. Halle & Vergnaud observed that the permitted consonant sequences were much freer in this context than elsewhere in the word. English displays final sequences four segments long in words like *sixths*. The sequence /ksθs/ does not occur word-internally. In order to express such skewing Halle & Vergnaud proposed an extra-rhymal constituent called the appendix. Its occurrence was limited to word margins and in this way they could offer a more elegant statement of consonantal distribution. Charette used the appendix to explain the apparent violations to prosodic government. She assumed that any final consonant occurred in the appendix position, which is to say, outside the rhyme. If the /p/ in *keep* was not in the rhyme then no violation of prosodic government had taken place. This is seen in (5) below:

(5)



Now the rhyme only contains the long vowel, [i:], and prosodic government is maintained. In the case of *kept*, only the [t] would be syllabified into the appendix. The stem-final [p] must go into the preceding rhyme and accordingly, shortening should take place.[9] Given that the appendix is independently motivated (its existence was postulated for reasons having nothing to do with the vowel-length facts), Charette's proposal went some way towards explaining this large class of apparent counter-examples to prosodic government.[10] Now one could distinguish the 'well-behaved' languages from the problematic ones by claiming that (i) the occurrence of the appendix in a language was a parameter and that (ii) English, French and Arabic had appendices, while Turkish and Yawelmani did not. In the latter cases single final consonants would, of necessity, be syllabified into the rhyme and thus the effects of prosodic government, to wit vowel shortening, must be observed in such cases.

This was the state of our understanding of the closed syllable vowel shortening phenomenon in the mid 80s. Charette's account seemed to offer a satisfactory explanation for a considerable proportion of vowel length alternations. By setting one parameter – the presence *vs.* the

absence of an appendix – one could distinguish languages which shortened before final single consonants from those that did not. But in fact, with the advent of a fuller theory of phonological government, major fissures began to appear in this analysis. To see why, I shall briefly outline the theory of phonological government as it first appeared at that time.[11]

2 Government Phonology and the role of empty nuclei

Phonological government is a theory about phonological strings, in the first instance. It defines under what conditions two phonological positions may be viewed as adjacent. Phonological positions are subject to the Licensing Principle given below:

(6) *Licensing Principle*
All phonological positions save one must be licensed within a domain. The unlicensed position is the head of this domain.

Government is one form of licensing. In order for A to govern B two conditions must be satisfied. These are given in (7):

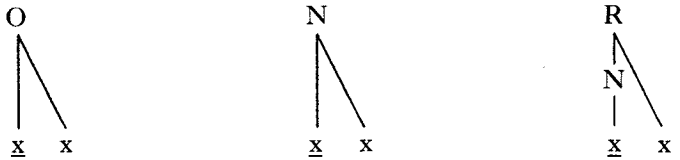
(7) *Strict Adjacency*
No position must intervene between governor and governee
Strict Directionality
Directionality of government at the skeletal level is universal (not parameterised as in syntax)

Two types of government follow from these definitions:

(8) Constituent government – Direction: left to right
Transconstituent government – Direction: right to left

We stipulate that there are three syllabic constituents: Onset (O), Nucleus (N) and Rhyme (R). Further, N is always the head (left branch) of R. The theory allows for the following maximal expansion of these categories:

(9)

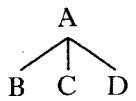


Given (7) above we can now derive the following theorem:

(10) *Binary Theorem*
All syllabic constituents are maximally binary

We can prove this inductively. Consider a ternary constituent:

(11)



The constituent A cannot contain a head which does not violate some aspect of (7). If either B or D is chosen as the head, then strict locality is violated. Choosing C will result in a strict directionality violation. Any constituent containing more than three members will contain the illicit configuration (11). Therefore, constituents are maximally binary. It can now be seen that prosodic government is just an instantiation of the Binary Theorem.

Within a constituent, heads require certain types of segments in order to fulfil their government requirements. In general governors are charmed[12] while governees are charmless. Oversimplifying considerably, low vowels and tense (ATR) vowels are (positively) charmed while lax non-low vowels are charmless. Governed members of nuclei must not only be charmless but also simplex. That is, they must be composed of a single charmless element. From this it follows that the off-glides of heavy diphthongs are limited to I° [ɪ], U° [ʊ] and v° [ɤ]. Given the asymmetric nature of government, if the sequence x_1 x_2 is a well-formed branching nucleus, then the sequence x_2 x_1 cannot be a well-formed branching nucleus. Thus, [aɪ] is a heavy diphthong but [ɪa] may not be so analysed. In fact, [ɪa] is a light diphthong with both segments being attached to the nuclear head position.

Non-nuclear segments are organised in similar fashion. Non-neutral stops and fricatives are governors.[13] In this instance charm theory gives results somewhat similar to a classification based on the sonority hierarchy although these approaches differ in some important respects. As above, the asymmetric nature of government is in evidence. [pl] is a possible branching onset with [p] occupying the governing position. [lp] is not a possible onset.[14] It is, as we shall see, a possible transconstituent sequence.

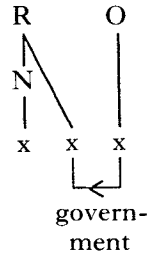
Constituent asymmetries of the type discussed above are important cues for establishing the existence of governing relations. These asymmetries are found in all syllabic domains. The question now arises if they may be found in adjacent positions that are not contained within a single constituent. When a closed syllable is followed by an onset we get a transconstituent sequence. A governing relation manifests itself by the severe phonotactic constraints that are found in this context. The definition of transconstituent government given in (8) above states that in these contexts government goes from right to left. This implies that transconstituent sequences should look like 'mirror images' of well-formed branching onsets. By and large this is true.[15] This implication only goes in one direction. It is *not* the case that the mirror image of any well-formed transconstituent sequence is a well-formed branching onset. This is due to the more stringent segmental requirements of governor and governee within constituents. In transconstituent sequences a governee must still be charmless.[16] Governors may or may not be charmed. If they are charmless, however, they must be more complex (contain more elements) than the segments they govern. [r] and [l] are both charmless.

However only [l] may govern [r] because of their segmental content, which is shown below:

(12) [r] R°
[l] R°.ʔ°

To sum up, the non-nuclear transconstituent governing domain is shown below:

(13)



The theory of phonological government is fraught with implications for the analysis of closed syllable shortening that was discussed above. In general any phenomenon that is sensitive to syllable weight should display the phonotactic constraints associated with transconstituent government. With this in mind, let us return to the data in (1).

Let us begin with the Turkish data. If what we are observing is indeed closed syllable shortening then we expect to find licit transconstituent sequences following the shortened vowel. Of course, if the vowel is word-final, then no such sequences occur and, presumably, all is as it should be. Such is the case for *merak* and *meraktan* (from *mera:k* and *mera:ktan*, respectively). The former shortening environment contains a single final consonant, *k*. The latter form has an internal closed syllable yielding the well-formed transconstituent sequent *kt* (cf. English *actor*, *doctor*, etc.). The plural form, *meraklar*, does not fare so well. If *a* has been shortened due to a closed syllable effect, then the syllable must be closed by a *k*. The following onset contains an *l*, yielding the transconstituent sequence *kl*. But this is impossible in government theory. A following onset is required to govern a preceding rhymal consonant. An *l* could never govern *k*, given their respective charm values. *k* may govern *l* and indeed it does, yielding Turkish transconstituent sequences *-lk-* as in *belki* 'perhaps', *halk* 'people, populace'. In fact, there are no phonotactic effects that are manifested by any of the three Turkish suffixes illustrated in (1). That is, *any* stem-final consonant may appear before *-dan*, or *-lar*.[17] This behaviour is not at all characteristic of transconstituent domains.

A similar problem crops up in the Yawelmani data. Shortening takes place before single final consonants as in Turkish. Likewise in word-internal position the putative closed syllable shortening results in bizarre transconstituent sequences. A sample of these sequences follow:

(14)

ph	sa**ph**in		pn	sa**pn**it		bh	go**bh**in
bn	go**bn**it		ml	ʔa**ml**al		xl	mo**xl**ol

In no case may the consonant in the governing position to the right govern the preceding position. Once again there do not appear to be any phonotactic constraints on the consonantal sequences following the word-internal shortened vowel.

Clearly something is amiss here. We need an explanation for the failure of the word-internal sequences in Turkish and Yawelmani to manifest the normal signs of a transconstituent domain (13). Furthermore, when this situation is compared with that of the 'ill-behaved' languages like English or French an important difference emerges. Instances of English or French word-internal shortening invariably involve well-formed trans-constituent sequences. Some examples follow:

(15)

Christ	st	Chri**st**ian			wild	ld	wi**ld**erness		
keep	pt	ke**pt**			heal	lθ	hea**lth**y		
ve:r	rt	ve**rt**	'green'		do:r	rm	do**rm**ez	'sleep'	
pe:r	rd	pe**rd**ez	'lose'		se:r	rv	se**rv**ir	'serve'	

These sequences differ markedly from those found in Turkish and Yawelmani. They are much more limited here and without exception form transconstituent sequences where right-to-left government is manifest. Note further that English and French may create secondary clusters as the result of syncope. In such cases a wide variety of new clusters may be created, many of which are not possible transconstituent sequences. Significantly, no shortening occurs in these cases:

(16)

trifle		f'l	tri**f'l**ing
(s')ent[e:]te 'persists in'		t'm	ent[e:]**t'm**ent 'persistence'

One final point needs to be made about Turkish and Yawelmani. Both of these languages have vowel ~ ø alternations. In both cases, the vowelless form provides a context for shortening:

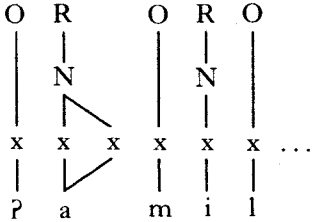
(17) *Yawelmani* ʔamlal ʔa:miltaw ʔa:milka

This behaviour is in sharp contrast to English, French or Arabic (to be discussed later) where such contexts do not result in shortening (cf. (16)). The vowel-shortening story provided by prosodic government along with the use of the appendix leaves a good number of unanswered questions concerning the facts described above. On the other hand, the theory of government is quite unequivocal in its handling of the word-internal Turkish and Yawelmani shortening cases: these could not be due to closed syllable effects for the reasons discussed above. Furthermore, KLV (this volume) posit a projection principle for phonology which drastically restricts the possibilities of resyllabification. If Turkish and Yawelmani are derived by syncope, contrary to common belief, an empty nucleus should remain and no closed syllable would be created.[18] In sum, if government theory is close to being correct, then the Turkish and Yawelmani shortenings have nothing to do with closed syllables![19]

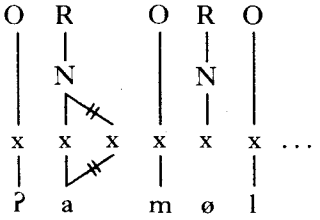
It is now reasonable to ask what the relevant context is for Turkish and

Yawelmani shortening. A clue is provided by the interaction between vowel ~ ø alternations and shortening. Given the projection principle, we must assume that the syllable structure of such forms is constant. To illustrate this let us look at the forms [ʔamlal] and [ʔa:miltaw]:

(18) a. [ʔa:miltaw]



b. [ʔamølal]



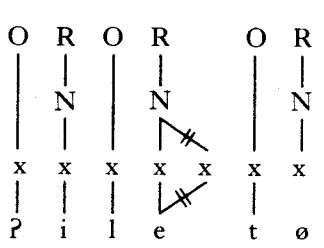
What distinguishes (18a), where vowel length is retained, from (18b), where it is not, is the nature of the following nucleus. In (18a) the vowel [i] follows the long vowel whereas in (18b) the nucleus following the long vowel is empty. Suppose then that this is the crucial distinction. We can now reformulate the shortening context as follows:

(19) A long vowel shortens when the following nucleus is empty

In a later section I will refine this context. I will also explain under what conditions an empty nucleus surfaces in these forms. For the moment let us assume that the empty/non-empty nucleus is the relevant distinction. This allows us to account for the cases where long vowels shorten word-internally in Yawelmani. It also explains the total lack of phonotactic constraints holding between the members of the apparent consonant clusters that following such shortening. No constraints should be expected since the consonants are in reality separated by a nucleus. Since this is not an environment for government we expect no interaction between the consonants in question.

This is all very well for the word-internal cases but we now must explain the shortening that occurs before single final consonants in both Turkish and Yawelmani. If the trigger for shortening is the presence of an empty nucleus following the long vowel, then it might be interesting to extend this notion to the word-final cases. The implications of this are that in Turkish and Yawelmani all domain-final consonants are followed by an empty nucleus. In other words, the final consonant of a domain is in the onset. To illustrate, let us take one case of Yawelmani final shortening:

(20) [ʔilet] < /ʔile:t/



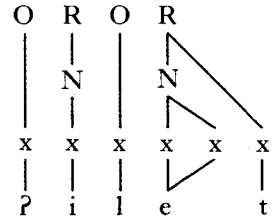
The structure in (20) satisfies the condition (19) and the vowel is observed to shorten. In sum, we can unify the contexts of this shortening if we assume that (20) expresses the relevant context and that domain-final consonants occupy the onset position. The latter assumption implies that these onsets are followed by an empty nucleus since every onset must be licensed by a following nucleus. Furthermore, (19) is not a principle of UG but rather a property that may or may not be present in a given phonological system. Turkish and Yawelmani (among others, of course) possess this property but English, French and Arabic do not. So some languages display shortening before empty nuclei and others do not. But what about the notion of domain-final consonants occurring in onset positions? Is this a parameter like (19), or is it part of the theory of syllable structure contained within UG? If the latter proves to be the case, then it should be as much a property of English, French and Arabic as it is of Turkish and Yawelmani. Let us pursue this idea.

First, we need a way to formulate the principle that domain-final consonants do not occur in rhymal positions. I will call this principle 'coda licensing', where 'coda' is shorthand for a post-nuclear rhymal position:

(21) *Coda licensing principle*
Post-nuclear rhymal positions must be licensed by a following onset

Let us apply (21) to a case like [ʔilet] (< /ʔile:t/). The normal assumption concerning the syllabification of this form is shown below:

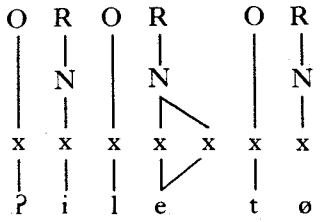
(22)



The standard solution would have *t* occupying the coda position of the final syllable. Since this syllable is closed, closed syllable shortening would apply, yielding the observed form. The structure in (22) violates the coda licensing principle. There is no following onset to license the occurrence of the rhymal *t*. If the position containing the *t* cannot occur in the rhyme, then it must be syllabified into an onset position. In turn, this onset must

be followed by a nucleus for the reasons stated above. This yields the structure in (23):

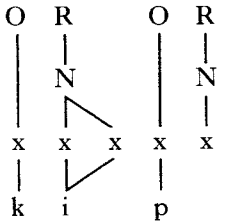
(23)



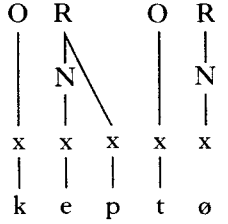
Crucially, the branching nucleus is followed by an empty one which triggers the shortening effect. We have thus provided an alternative account to the closed syllable analysis which does not involve violating government restrictions on transconstituent sequences. If coda licensing is a principle of UG, we should be able to detect its effects in languages other than Turkish and Yawelmani. In fact, coda licensing is making the claim that no domain can terminate in a closed syllable. This means that we should be able to see differences between true closed syllable effects in word-internal position where the rhymal consonant may be licensed by a following onset and spurious closed syllables which contain a domain-final consonant which should be syllabified as an onset.

Notice that English and French both lack long vowels in word-internal closed syllables. Unlike the Turkish and Yawelmani cases, the following clusters are well-formed from a government point of view (see the data in (15)). Assuming that the absence of long vowels in these contexts is a genuine closed syllable effect, we then predict that long vowels should freely occur before single domain-final consonants precisely because these syllables are *not* closed. Coda licensing requires the structures show below for the English forms *keep* and *kept*:

(24) a.



b.



In (24a) the domain-final *p* is in onset position. There is no licensing onset that would allow syllabification into the rhyme. Since the long vowel of *keep* is in an open syllable, there is no necessity to have a short vowel in this position. The situation is different in (24b). Here the stem-final *p* is followed by an onset occupied by *t*. This satisfies the coda licensing condition (21) and the *p* of the stem is syllabified into the rhyme. Since a long vowel may not occur in a branching rhyme, the vowel is short in this form. Principle (21) has correctly predicted the asymmetry of so-called closed syllable effects in English.[20] In a similar manner, the same asymmetry between final and non-final 'closed syllables' is explained for

French, Arabic and Wolof. Indeed (21) predicts that true closed syllable effects should *never* be observed before single domain-final consonants.

We can now summarise the difference between languages like Turkish and Yawelmani, on the one hand, and English, French, Arabic and Wolof, on the other. In the former case, the observed shortening phenomenon is not due to closed syllable effects but rather to the context of a following empty nucleus. This sensitivity is a parameterised variable in UG and as such it is not omnipresent in phonological systems. Some languages display it, others do not. The shortening occurring in English is a manifestation of the properties of syllable structure. This restriction forms part of UG and should be present in every language that contains branching nuclei. Indeed, it should be present in Yawelmani and Turkish as well as English. It is ironic that the examples in (1) originally designed to illustrate closed syllable effects do not contain a single example of a closed syllable. This is not to say that Turkish and Yawelmani do not have closed syllables. Rather, the vowel-length alternations that are presented in (1) are not examples of such a syllable type.

3 Licensing empty nuclei

Let us now consider the status of the empty nuclei which are so crucial for the Turkish and Yawelmani cases. Obviously, empty nuclei cannot be used as a 'phonological seasoning' to be sprinkled over phonological representations whenever their presence is required. In fact, empty nuclei will be realised phonetically unless some very special conditions are satisfied. The phonetic realisation of empty nuclei offers concrete evidence of their presence in the string. It remains to lay out the conditions under which we expect to 'hear' these nuclei if they are indeed present. Government phonology contains its own version of the Empty Category Principle (ECP):

(25) *Empty Category Principle*
A properly governed empty nucleus has no phonetic realisation

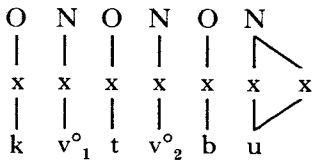
In other words, an empty nucleus will be inaudible if it is properly governed and audible otherwise. The specific realisation will vary according to the language but in the unmarked case we expect to hear the cold vowel, $v^{\circ} = [\text{ɪ}]$. I must now offer a definition of 'proper government', a condition that must be satisfied if an empty nucleus is to have no phonetic content:

(26) A nuclear position α properly governs a nuclear position β iff
a. α is adjacent to β on its projection
b. α is not itself licensed
c. No governing domain separates α from β

Proper government is a relationship between two nuclei. A nucleus is properly governed if it is immediately preceded or followed by another

nucleus (the governor) in the nuclear projection which is not itself properly governed or licensed in some other way and if no governing domain intervenes between governor and governee.[21] To illustrate proper government and how it results in vowel ~ ø alternations, consider the following example from Moroccan Arabic:[22]

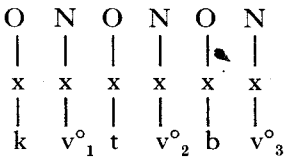
(27) O N O N O N [kɨtbu:] '(they) write'



Starting from the right edge we find an unlicensed nucleus, [u:]. The immediately preceding nucleus contains a cold vowel, $v°_2$. No governing domain intervenes and so $v°_2$ is properly governed by the final nucleus. The ECP is in effect, so $v°_2$ receives no phonetic content.

What about $v°_1$? The potential proper governor for $v°_1$ is $v°_2$. But this nucleus is itself licensed by [u:] and so it cannot serve as a proper governor. Therefore $v°_1$ is not properly governed and it surfaces as the cold vowel, [ɨ]. Let us now look at the corresponding singular form. This is like the plural form but without the suffix [-u:]:

(28) O N O N O N [ktɨb] '(he) writes'



Working again from right to left, consider the final nucleus, $v°_3$. This nucleus is not phonetically realised, which means that it must be licensed in some way. Note that there is no following nucleus, so whatever licensing is involved it cannot be proper government; there is no proper governor. Suppose we propose that domain-final position is a potentially licensed position. What this means is that any empty nucleus occurring in that position is licensed. This makes a specific prediction. Since Arabic would license final empty nuclei and since non-licensed empty nuclei are realised as [ɨ], it must be the case that Arabic contains no domain-final [ɨ] since this would be licensed and hence inaudible according to the ECP. This prediction turns out to be correct.

We must now slightly modify the ECP to allow for licensing cases other than proper government:

(29) *Empty Category Principle* (revised version)
- i. A licensed empty nucleus has no phonetic realisation
- ii. An empty nucleus is licensed if (a) it is properly governed or (b) if it is domain final in languages which license domain-final empty nuclei

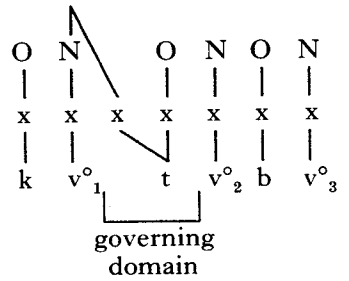
As can be seen from (29), the licensing of final empty nuclei is treated as a parameter. Some languages have this property; others do not. Languages

that do not license final empty nuclei cannot have domain-final consonants. Coda licensing requires that such consonants be in onset position, necessitating a following (empty) nucleus. Since this nucleus is neither properly governed[23] nor is it licensed by virtue of being in final position, it must be realised phonetically in order to avoid an ECP violation. So 'CV' type languages such as Desano, Lingala or Vata do not license final empty nuclei. Traditionally languages of this sort have been described as 'open syllable languages' and the absence of final consonants in these languages is thought to be because they lack closed syllables, the implication being that a domain-final consonant closes the preceding syllable. Below I will present evidence to show that the absence of final consonants is independent of whether or not a language contains closed syllables.

Continuing our discussion of (28) we see that $v°_3$ is not a proper governor because it itself is licensed. Therefore, $v°_2$, which is neither properly governed nor in domain-final position, must be realised phonetically. Now consider $v°_1$. The following nucleus is $v°_2$ and we have just seen that it is not licensed. It is therefore an eligible proper governor. It exercises this role over $v°_1$ which is now licensed and hence inaudible as per the ECP. This yields the correct output, as shown in (28).

Finally, let us consider a causative verb. The syllable structure of this verb differs from what we have seen up to now in that the second literal is geminated:

(30) [kɨttɨb] '(he) causes to write'



The final empty nucleus is licensed and accordingly is neither audible nor a potential governor for $v°_2$. Since $v°_2$ is not licensed it will be pronounced. One might expect that $v°_2$ could serve as a governor for $v°_1$ since the former is not licensed. However, a (transconstituent) governing domain intervenes between $v°_2$ and $v°_1$ which blocks proper government. Therefore, both $v°_2$ and $v°_1$ are pronounced as indicated in the phonetic transcription.

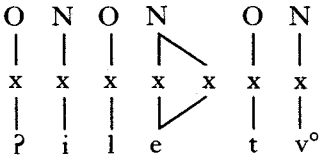
With this analysis in mind, let us turn to the Yawelmani data. Considerations of government required positing an empty nucleus word-internally that would have the double effect of separating consonants that could not form a cluster and of triggering the Yawelmani shortening, which I claim is not a closed syllable effect. Using the ECP as in the Arabic example I will now try to motivate the presence of these empty nuclei. I

will show that the Yawelmani facts are virtually identical to those of Arabic. To begin with let us consider the following forms:

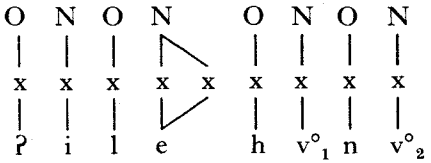
(31) ʔilet ʔile:hin

This forms will have the structures shown below:[24]

(32) a.



b.



Consider first (32a). Coda licensing requires that the word-final *t* be followed by an empty nucleus. This triggers the Yawelmani shortening, yielding [ʔilet]. In (32b) the stem is followed by the suffix *-hin*. I have assumed that the vowel appearing in this suffix is an empty nucleus, although nothing in the data at hand depends crucially on this. Like Arabic, Yawelmani licenses final empty nuclei. The nucleus $v°_1$ is not licensed and must be realised phonetically. The only difference between Yawelmani and Arabic is that in the latter system the cold vowel may be realised as such, i.e. as [ɨ], whereas in Yawelmani an epenthetic element I° must be added to yield [i]. There is no context that can trigger the Yawelmani shortening process and so the vowel remains long.

Let me digress briefly here to discuss the actual context for Yawelmani (and Turkish) shortening. My first formulation of this process appeared in (19) above. There I stated that the relevant context was a following empty nucleus. Notice that in (32b) the derivation starts with the long vowel /e:/ being followed by an empty nucleus. This nucleus gets spelled out as [i] in the course of the derivation once it is determined that it is unlicensed. This would appear to imply some sort of ordering between the spelling out of this nucleus and the application of the shortening process. It appears to imply that if there was another Yawelmani, call it Yawelmani*, identical in all respects save that it allowed the cold vowel to surface in its pristine form [ɨ], then shortening would take place. The idea would be that representationally, there is no difference between an audible and an inaudible cold vowel. Only relations of licensing would distinguish these cases. It is my belief that Yawelmani*'s shortening process would apply in exactly the same forms as does Yawelmani's. Whether or not the vowel surfaces as [i] or [ɨ] is irrelevant. In fact, this is more than mere speculation. As we shall see, Turkish is much like Yawelmani* and it

behaves exactly like Yawelmani with respect to shortening. We need to revise the triggering context for shortening in the following way:

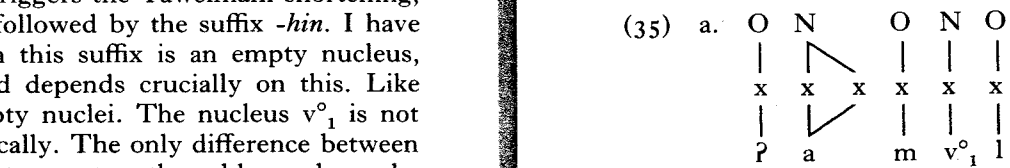
(33) A long vowel shortens when the following nucleus is *a licensed empty nucleus*

Returning to (32b) and applying (33) we see that the long vowel would not shorten, regardless of how the nucleus is realised. In any case it is not licensed and so the conditions for shortening as now defined are not satisfied. To conclude the Yawelmani discussion, I will discuss the data given below:

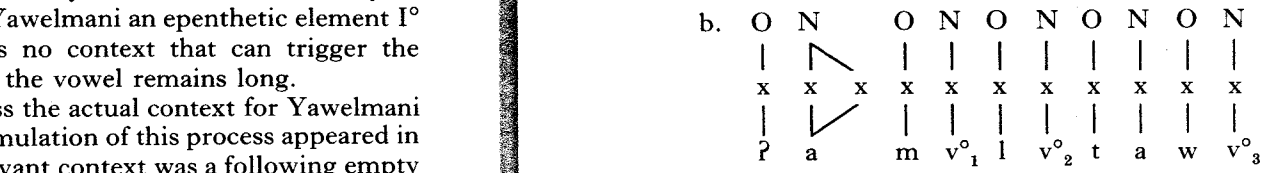
(34) ʔamlal ʔa:miltaw

The standard solution would portray the underlying form of this verb as /ʔa:ml/. A rule of epenthesis, crucially ordering with respect to shortening, would generate the forms shown above. Government phonology excludes clusters of the form *-ml-*. *l* cannot govern *m* as would be required for a transconstituent sequence and Yawelmani does not have branching onsets. Even if it did, a charmless segment like *m* cannot be the head of a branching constituent. We are therefore required to posit an empty nucleus between the *m* and the *l*. We are now in a position to provide the relevant structures for the forms of (34):

(35) a.



b.



We begin with (35a). The final empty nucleus is licensed. The preceding nucleus is not empty and thus does not fall under the ECP. The first empty nucleus $v°_1$ can be licensed by the following *a*. It is inaudible and as a licensed empty nucleus it triggers the shortening of the previous vowel. The derived form is [ʔamlal], which corresponds to the datum in (34).

In (35b) the suffix *-taw* is appended to the stem. As before, the final nucleus is licensed and the preceding one is not empty. $v°_2$ is licensed by the following nucleus. Since $v°_2$ is licensed it cannot serve as a governor for $v°_1$. This latter nucleus is eventually spelled out as [i]. Since the long vowel [a:] is not followed by a licensed empty nucleus, it does not shorten and we derive the correct form [ʔa:miltaw]. The remaining Yawelmani examples are amenable to the same analysis and the results appear to be

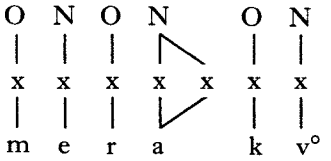
correct for these cases and for many others not discussed here. What should be emphasised is that Yawelmani and Arabic display nearly identical behaviour with respect to the distribution of empty nuclei. Only the realisation of unlicensed empty nuclei distinguishes these cases.

Turkish, like Arabic, does permit empty nuclei to be realised. The vowel [ɨ] is part of the Turkish nuclear inventory. Thus, as in Arabic, the cold vowel will be inaudible when licensed and realised as [ɨ] otherwise. Turkish vowel shortening takes place in identical circumstances to that of Yawelmani, *viz.* before licensed empty nuclei. With this in mind we proceed to some examples:

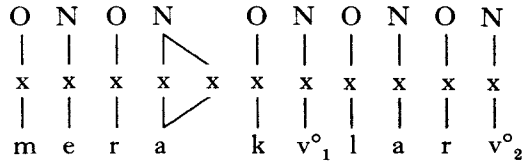
(36) merak mera:kɨ meraklar

The stem for this form is *mera:k*. Let us first consider the forms [merak] and [meraklar]:

(37) a.



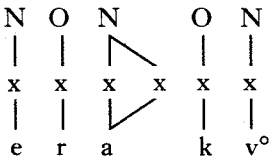
b.



(37a) represents the final empty nucleus case. As in Arabic, Turkish licenses final empty nuclei and so shortening takes place in this example. Likewise in (37b) we have a nucleus $v°_1$ which is licensed by the following *a* and this serves as a trigger to shorten the preceding long vowel [a:].[25]

Let us now take a look at the possessive form [mera:kɨ]. Given its phonetic transcription, at first glance all appears to be in order. However, when we look at the structure (38), which one presumes underlies this form, things are far from clear:

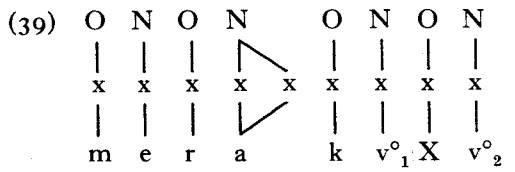
(38)



The problem is that (38) is identical to (37a). In the case of (37a) the final empty nucleus is licensed and does not appear phonetically as [merak]. This licensed empty nucleus triggers shortening, as we have seen above. In (38) we have the form [mera:kɨ]. The final vowel is obviously not licensed, because it has phonetic content and so there is no shortening. But Turkish licenses final empty nuclei and, as we have stated above, the cold

vowel is realised as [ɨ] in Turkish when not licensed. This leads us to posit the structure in (38), but this is precisely the same structure as in (37a), which underlies [merak]. Why is the final empty nucleus licensed in one case but not in the other? Clearly one of these structures must be wrong. Since we have every reason to believe that Turkish licenses final empty nuclei – Turkish abounds with consonant-final forms and coda licensing requires that these be followed by an empty nucleus – we can only conclude that the structure (38) cannot underlie [mera:kɨ]. Otherwise there is something desperately wrong with the theory of coda licensing. If this theory is correct then the audible cold vowel in [mera:kɨ] cannot be domain final; otherwise it would be licensed and not audible. Suppose that the theory is correct. There could be a following consonant that is not phonetically realised for some reason or another, which would strip v° of its licensed position. If it is not licensed then of course it will be pronounced and we can derive [mera:kɨ]. The structure I have in mind would look something like the following:

(39)



Here the possessive suffix is not simply -v° but rather -v°X where 'X' represents some mysterious consonant which is not realised phonetically for one reason or another. Notice that now $v°_1$ is not in a licensed position and will appear phonetically in this form. The mystery consonant 'X' must of course be followed by another empty nucleus, $v°_2$. Unlike $v°_1$ this nucleus is licensed and thus inaudible. The structure (39) is clearly different from (38) which accounts for the observed differences between [merak] and [mera:kɨ].

The above account will indeed handle the possessive forms, but it is far from complete. For one thing, the insertion of a mystery consonant 'X' has all the appearance of an ad hoc measure whose only purpose is to save coda licensing from a difficult situation. Coda licensing makes the claim that the apparently final vowel in the possessive form [mera:kɨ] could not possibly be final. Is there any independent motivation for the presence of a mystery consonant that follows $v°_1$ in (39)? Indeed there is evidence for this mystery consonant, precisely as predicted by coda licensing.[26] The following quotation from Underhill (1976: 90) provides the evidence:

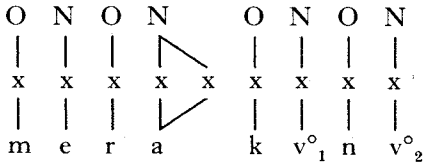
> You will have noticed by this time that the pronoun *o* adds an *n* before any case suffix (*onu*, *ondan*) and the plural suffix (*onlar*). That is, all case forms and the plural are built on an 'oblique stem' *on-*, although the final *n* is missing when the word stands alone.

Underhill offers the following forms to illustrate his point (1976: 90):

(40)	baba	'father'	ev	'house'
	babasi	'his father'	evi	'his house'
	babasina	'to his father'	evine	'to his house'
	babasindan	'from his father'	evinden	'from his house'

To these forms should be added *babaya* 'to the father', *babadan* 'from the father', *eve* 'to the house', *evden* 'from the house'. These examples show that the *n* found in the combinations stem + POSS + CASE/PLURAL is neither part of the stem nor part of the case/plural suffixes. As Underhill suggests, it must be considered to be part of the possessive suffix, being deleted only in word-final position. The above provides the needed independent support for the mystery consonant 'X'. Further, it identifies this consonant as being an [n].[27] We can now replace 'X' by *n* in (39), yielding (41):

(41)



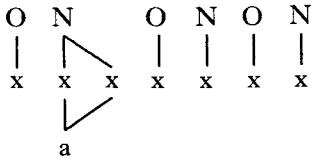
Now we see that $v°_1$ is not in a licensed position. It is not domain-final and there is no available proper governor. The ECP requires that it be realised phonetically, which it is.

In sum, our conclusion that Turkish vowel shortening is triggered by a following licensed nucleus seems well supported by the data. I conclude that neither Turkish nor Yawelmani shortening is related to any closed syllable effects. Languages which do not shorten before single final consonants, such as English, French, Arabic and Wolof, do display genuine closed syllable shortening effects. Typically these are not manifested by phonological processes but rather are reflected distributionally.[28] Let us consider the case of Moroccan Arabic (MA). Like any other language MA does not contain branching nuclei within branching rhymes. Words ending in a long vowel followed by a single consonant abound, but, given coda licensing, these forms are now easily explained. The following example will serve to illustrate the exclusion of long vowels from closed syllables in MA.

The MA imperfect verb template consists of three open syllables. MA has the property of allowing only empty short vowels, i.e. only the cold vowel may be associated to a single nuclear position. As has been stated above, final empty nuclei are licensed in MA and an unlicensed cold vowel surfaces as such. It follows then that [ɨ] never occurs in word-final position. Just as in Turkish, this holds for MA. Triliteral verbs are then typically realised as *CCɨC*, as we saw in (28) above. The active participle

for such verbs shows a similar template, except that the first nucleus branches and is associated with the segment *a*:

(42) *MA Active Participle Template*

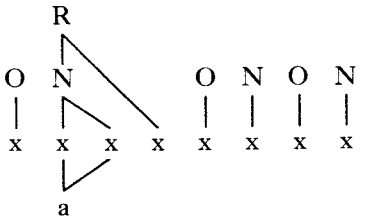


Inserting stems such as |ktb| 'write' or |dxl| 'enter' yields [katɨb] and [daxɨl], respectively. Insertion of the branching nucleus in this context causes no problem since the syllable is open. No constraints are violated. What is of interest here are similar forms involving quadriliteral or causative stems. The template is that given in (30) above:

(43)

(43) represents an imperfective form. Notice that the initial syllable is closed. Recall that the active participle of triliteral stems differed from the imperfective form in that the former contained a long [a:] where the latter had a non-branching nucleus. Suppose we try to utilise the same strategy for forming active participles of quadriliteral stems. This would result in the following structure:

(44)

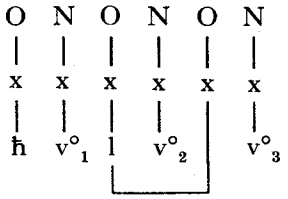


In other words, we would expect an imperfective form like [hɨrrɨs] 'breaks' to have an active participle [ha:rrɨs]. However, it should be clear that the structure (44) is ill-formed. It contains a branching nucleus within a branching rhyme. We must predict then that (44) could not be the active participle structure for quadriliterals and causatives in MA. This turns out to be correct. These stems form their active participle by prefixation and not by mutating the initial nucleus. The active participle for the verb 'break' is not *[ha:rrɨs] but rather [mhɨrrɨs], formed by adding the prefix, *m*v°-. This entails no violation of government theory. The above facts gives us an indication that MA eschews long vowels in closed syllables, as we expect.

An interesting apparent counterexample is the active participle [ħa:ll] of

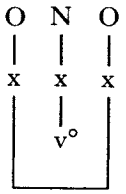
the verb 'open'. Here we find an apparent case of a long vowel followed by *two* domain-final consonants. This should be clearly impossible. In fact, the aberrancy of this form is only an illusion. This becomes clear when we realise that biliteral stems such as |ħl| use the triliteral template. When we associate |ħl| to the template (28) we get the following:

(45)



Following the standard assumptions for Semitic template associations first proposed by McCarthy (1979), the segments of the stem are linked to the template in a left-to-right fashion. Since there is one more non-nuclear position than there are consonants, the stem-final *l* is joined to two positions. However, these positions are separated by an empty nucleus since this template consists of three open syllables, as we saw in (28). Normally we would expect v°_2 to be realised since it cannot be properly governed. This is the case in a form like [kt‡b]. However (45) yields [ħ‡ll] and not *[ħl‡l]. If v°_2 is not realised it must have satisfied the ECP in some way. I will assume that any doubly-linked segment forms a governing domain. Further, I will add a proviso to the ECP that this governing domain licenses an empty nucleus contained within it. Schematically this means that

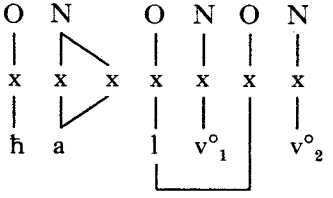
(46)



will meet the ECP requirements for proper government.[29] The empty nucleus contained within this structure is not realised. This is not to say that the nuclear position is deleted. As we will see anon, it is vital that this position remain.

If (45) represents the imperfective template, then we would expect the active participle to insert a long [a:] in the initial nucleus, just like any other triliteral stem. If this takes place we get (47):

(47)



As before, v°_1 will not be pronounced. It is embedded within a governing domain. There is no problem with the initial long vowel since it does not occur in a branching rhyme. Thus, the form [ħa:ll] in no way contradicts the general principles of syllable structure excluding long vowels in closed syllables. This form has no closed syllables. The form [ħa:ll] has what we may term a 'spurious geminate', i.e. one with an intervening nucleus. Spurious geminates do not show closed syllable effects whereas true geminates do. Comparing this form with a true geminate brings out the difference in behaviour:

(48)

Geminates	*true*	*spurious*
IMPERFECTIVE	h‡rr‡s	ħill
ACTIVE PARTICIPLE	mh‡rr‡s	ħa:ll

4 Coda licensing and syllable typology

I will now turn to a discussion of coda licensing as it concerns questions of syllabic typology. It has long been thought that languages differ along a closed/no closed syllable parameter.[30] We can contrast a language like Desano, which contains only open syllables, with a language like Arabic, which has both open and closed syllables. Desano is a typical 'CV' language with no consonant clusters and no word-final consonants. In this language sequences like *barba* and *bar* are ill-formed. Both are possible in Arabic. It has generally been assumed that *barba* and *bar* are excluded from 'CV' languages like Desano for the same reason, *viz.* such languages have no closed syllables and forms like *barba* and *bar* both contain this syllable type. The same reasoning would predict that both these forms are possible in Arabic. Again this assumes that both forms contain closed syllables. Leaving aside questions of branching onsets or nuclei, the above assumption predicts two types of languages. This prediction is schematised below:

(49)

Closed syllable	*No closed syllable*
CV	CV
CVCCV...	*CVCCV...
CVC#	*CVC#

If we accept coda licensing then the above typology is no longer tenable. The presence or absence of genuine word-internal consonant clusters is indeed due to the setting of the branching rhyme (closed syllable) parameter. However, this parameter is unrelated to the occurrence of domain-final consonants. This latter property is due to whether or not the system in question licenses final nuclei. If it does, the language will contain domain-final consonants. Otherwise all domains will end in

phonetically realised nuclei. This theory predicts the existence of four, rather than two, types of systems:

(50)

Final empty nuclei licensed	Branching rhymes: *yes*	Branching rhymes: *no*
yes	CVCCV ... CVC#	*CVCCV ... CVC#
no	CVCCV ... *CVC#	*CVCCV ... *CVC#

Language types found in the upper left and lower right cells of (50) correspond to the closed *vs.* no closed syllable types found in (49). Coda licensing predicts two additional types, however. This follows from the independence of the two parameters involved: branching rhymes and licensing of final empty nuclei. Concretely, coda licensing predicts that we should find languages that contain word-internal closed syllables but with no domain-final consonants. This corresponds to the lower left cell in (50). There should also be languages that contain no word-internal closed syllables but still have (phonetic) domain-final consonants. In fact both these predictions are borne out. Japanese does contain branching rhymes but does not license final empty nuclei. In traditional terms, it has word-internal closed syllables but contains the restriction that no word may end in a closed syllable.[31] A number of Gur languages of Burkina Faso and some Cameroonian languages appear to have the property of permitting no branching rhymes but allowing domain-final consonants. Thaïs da Silva reports that Krenak, an Amerindian language spoken in the state of Minas Gerais, is also of this syllabic type. More research is required on these languages.

Coda licensing also sheds some light on certain rather odd correlations that are expressed as below in rule-based systems:

(51) a. In languages in which shortening takes place before a single final consonant, no phonotactic constraints will hold in word-internal context where shortening is also observed

b. In languages in which shortening takes place before a single final consonant, vowel syncope will always feed shortening.

The generalisations of (51) are rather odd, to say the least. (51a) involves forms such as Yawelmani [ʔilet] < /ʔile:t/ and [saphin] < /sa:phin/. The former case illustrates shortening before a single final consonant. In the latter case, the standard analysis holds that the *a:* shortens as a result of it occurring in a syllable closed by *p* and followed by another syllable whose onset is *h*. (51a) states that there will be no phonotactic constraints involving the purported coda consonant and the following onset. Note that internal shortening in languages like English or French has very strong phonotactic constraints in this context, as was seen in (15) above.

(51b) states that in languages like Yawelmani and Turkish (unlike French and English), a 'syncope rule' will feed the shortening process. Consider a form like [ʔamlal] < /ʔa:milal/. If the loss of the vowel *i* is considered to be a syncope process,[32] then the standard view would have syncope creating a new environment for the application of shortening as a result of the 'creation' of a closed syllable [ʔa:m]. Notice that neither French nor English, which do not shorten before single final consonants, displays any sensitivity to syncope processes, as the data in (16) show.

Now why should these generalisations hold? Coda licensing provides a clear explanation. Any case of shortening taking place before a single final consonant is *not* a closed syllable phenomenon. It is rather a process triggered by a licensed empty nucleus. If shortening takes place word-internally then the putative clusters that are involved, i.e. the 'coda consonant' followed by the onset consonant, are spurious clusters. They are in reality separated by an empty nucleus. In such a context there is no particular governing relation that need hold between them. Since governing relations are the source of most phonotactic constraints, we do not expect to find these constraints in contexts where governing relations are absent. The so-called clusters are nothing more than sequences of onsets and there are generally no constraints on onset sequences. English and French shortenings are genuine manifestations of closed syllable phenomena. Since the rhymal consonant must be governed (and licensed) by the following onset, strict phonotactic constraints should be observed, as indeed they are.

(51b) is also a natural consequence of coda licensing. Given the projection principle in phonology (KLV this volume), syncope (or epenthesis) cases as seen in Yawelmani, Turkish, French and Arabic are manifestations of proper government. The syllable structure (specifically, the nucleus in question) is preserved. What changes is the licensed status of this nucleus according to the availability of a following proper governor. Proper government will then have two effects: the licensed empty nucleus is inaudible and the conditions for the special shortening case are met. Syncope (or epenthesis) occurring in languages which display true closed syllable shortening will have no effect. Since these events do not create closed syllables, there is no reason for shortening to take place in such situations.

It appears that the empirical record offers some support for the claim that the structures CVCCV... and CVC# do not belong to the same syllable type. This is exactly what the coda licensing principle would lead us to expect. To this point all the evidence adduced in favour of this contention involves nuclear branching. We should be able to find other processes that are sensitive to rhymal branching. We predict that such processes should treat the strings CVCCV... and CVC# differently. One likely candidate for such a test is stress assignment. Many stress systems are quantity sensitive (Hayes 1980). We would expect to find examples where a so-called closed syllable in domain-final position behaves as if it were open. Concretely, suppose we have a system which stresses the first

(i.e. leftmost) heavy syllable, otherwise stressing the initial syllable of the domain. This gives the patterns shown below:

(52) pártila sukénta málisu vukasárd

Now what stress pattern would we expect to find for a word like *kisilak* in this system? Given coda licensing, the predicted stress pattern is *kísilak*. This is because this form is required to have the following structure:

(53)

O	N	O	N	O	N	O	N
x	x	x	x	x	x	x	x
k	i	s	i	l	a	k	v°

As can be seen from (53), *kisilak* has no heavy syllable. In such cases stress occurs on the first syllable, yielding *kísilak*. This is all very well as an hypothetical example, but do such systems actually occur? Indeed they do. In such cases the apparently aberrant behaviour of the 'final closed' syllable is handled by treating the final consonant as EXTRAMETRICAL. In other words, the word-final consonant behaves as if it weren't there. It is interesting to note that coda licensing can provide a principled account for many, if not all, of the cases requiring extrametricality. The extrametricality facts give another indication that coda licensing is saying something deep about syllabic structure.

Coda licensing may provide explanations for other seemingly arbitrary facts about stress systems. Segundo (1989) discusses the Natal stress system of Brazilian Portuguese in some detail. She notes that Natal stress may be final, penultimate or antepenultimate. However, consonant-final forms (not counting plurals; see Segundo 1989 for details) can never have antepenultimate stress. Normally, the possibility of having such stress depends on the nature of the *penultimate* syllable (it must be light). Coda licensing explains why such an apparently non-local condition should hold in the Natal stress system. To place stress on the supposedly antepenultimate syllable of a word like *támaras* would in reality have the stress appearing on the syllable *preceding* the antepenultimate syllable. Coda licensing requires that the *s* be followed by an empty nucleus. This predicts possible stressing for this form as *tamáras* or *tamarás*. Thus, **támaras* is impossible in Natal for the same reason that **ámerica* is.

It is clear that a considerable amount of reanalysis is required to verify the predictions of coda licensing with respect to stress systems. The above discussion is meant to be suggestive and the results would indicate that this task may well be worth the effort. There are still many mysteries that surround the examples discussed in this article. The nature of the shortening process that is sensitive to following licensed empty nuclei needs to be understood. If we wish to exclude arbitrary processes from phonology, one of the primary goals of government phonology, then we should be able to derive this effect from some principle of grammar. This has yet to be accomplished. Proper government and the ECP involve some

unsolved problems. For example, many languages appear to resist proper government of domain-initial nuclei even when all the other conditions for such government are met. This appears to be the case in Yawelmani and Turkish. Tangale and Tonkawa are other systems where this attribute may be present. It is not a universal property of phonological systems, however. French, Arabic and Polish can and do have properly governed empty nuclei in the first syllable of a domain. On the other hand, languages of the Yawelmani type are frequent enough for the initial nucleus restriction to be more than some minor irregularity. It seems that there are other parameters involved in this story. It is to be hoped that these problems, far from stifling interest in this enterprise, will stimulate research in this area. I remain convinced that this area of research holds many more surprising and rewarding solutions.

NOTES

* I am grateful to Jean Lowenstamm, Jean-Roger Vergnaud, Mohand Guerssel, Monik Charette and two anonymous *Phonology* readers for comments on earlier drafts of this article. I alone am responsible for any errors found herein.

[1] See Kaye & Lowenstamm (1981, 1984), Lowenstamm & Kaye (1985), Kaye (1985), Kaye *et al.* (henceforth KLV) (1985, 1989, this volume).

[2] The well-foundedness of this enterprise has hardly met with universal approval. For example, Bromberger & Halle (1989) explicitly reject this possibility.

[3] The Turkish and Yawelmani data are from Halle & Clements (1983: 137, 139; 153, 155).

[4] It should be noted that vowel length is not a property of native Turkish vocabulary. All the Turkish stems in (1), as well as others manifesting vowel length, are of Arabic or Persian origin. Furthermore, the vowel length of these loan words is not maintained in all Turkish dialects.

[5] This rule does not represent the 'state of the art' in non-linear rule formulation. Modern theories would characterise the shortening of the vowel by the deletion of a skeletal point of 'V-slot'. Nothing essential hinges on this simplification.

[6] Prosodic government was first discussed in Kaye & Lowenstamm's GLOW paper of 1982. This paper was published as Lowenstamm & Kaye (1985).

[7] In Quebec French *all* the exceptional cases occur in this context. English does contain branching nuclei before final consonant clusters but only if these are coronal clusters, e.g. *child, mount, weird.* Notice that shortening still takes place when these coronal clusters are non-final as in *children, wilderness, Christian*, etc.

[8] Cf. Kaye (1990). MA does have apparent cases of CVVCC but, as we shall see below, the final consonant sequences are spurious.

[9] In fact recent work on lexical representations and lexical access by Vergnaud and the author indicate that *keep* and *kept* are separate lexical entries (cf. Kaye & Vergnaud 1989). English thus displays the closed syllable shortening effects in terms of distributional restrictions rather than through alternations. These latter effects can be seen in cases of metrical lengthening in languages such as Italian and Norwegian.

[10] Charette's proposal for the appendix differed somewhat from that of Halle & Vergnaud in that she proposed that single final consonants were syllabified into the appendix and not into the rhyme. This meant that the appendix was not limited to coronal consonants as Halle & Vergnaud proposed.

[11] For a more complete description of the theory of phonological government see Charette (1988), KLV (this volume), Kaye (1990).

[12] The theory of charm forms part of the theory of segmental representations. See KLV (1985) for a discussion of charm theory with respect to nuclear segments.

[13] Non-neutral segments involve a laryngeal element. This may be a high tone H^- (stiff vocal folds) or a low tone L^- (slack vocal folds), perhaps in combination with the noise element h° (expanded glottis).

[14] Claims made about possible onsets are routinely contradicted in the phonological literature. Much of the controversy involves the (often tacit) assumption that any sequence of initial consonants up to the first nucleus of the word constitute an onset. Arguments against this position are found in KLV (this volume).

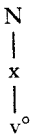
[15] See KLV (this volume) for a discussion of the differences between branching onsets and transconstituent sequences of the type under discussion.

[16] One example of this departure from complete mirror-image relations is the asymmetry between the licit transconstituent sequence $[p^\circ t^-]$ and the ill-formedness of the putative onset $*[t^- p^\circ]$. Governed members of constituents are not only charm neutral but also simplex.

It should be noted that the theory requires two types of *p*'s and *k*'s for English: a charmed version which may be a constituent head as in *pray* and *clay* and a neutral version for *adopt* and *doctor*. Indeed the governed stops in the last two forms are unreleased and lax.

[17] One can observe that there are changes in voicing taking place in both the stem-final consonants and the suffix-initial *d*. These will be discussed below.

[18] Following the proposal set out in KLV (1985) an 'empty nucleus' is one that contains only the cold vowel, v°, i.e. it has the structure:

(i)


[19] It is worth considering whether the loss of a nuclear point in the course of a phonological derivation is in violation of the projection principle. Notice that such a loss does not involve changing the nature of governing relations. The head of the nucleus remains the head (on the assumption that it is the governed member of the nucleus that is lost). No formerly governed position is elevated to the status of a governor.

[20] What (21) does not do is explain the absence of shortening when branching nuclei are followed by coronal sequences as in *child*, *pint*, *wield*, etc. For the moment I have no ready answer for this problem.

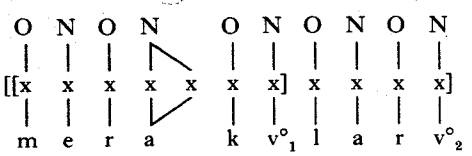
[21] Until now all clear cases of proper government have been invariably from right to left. Indeed Jean Lowenstamm has suggested that proper government may only proceed in a leftward direction. Note that other internuclear relations such as harmony and stress are parameterised for direction. I will remain open on this question for the time being.

[22] The analysis is presented in detail in Kaye (1990). For an analysis of French schwa using proper government see Charette (1988). An extension of this analysis to Russian yers is found in Kaye (forthcoming).

[23] This assumes that proper government is invariably right to left. This assumption may not be correct.

[24] Various other processes affect the quality of the vowels in Yawelmani. I will not consider these processes here.

[25] There is a more likely story here. If the plural suffix *-lar* is cyclic, then the structure would be as follows:

(i)


In this case v°_1 is in domain-final (i.e. cycle-final) position. Final licensing would be applicable here rather than proper government. The net effect would remain. The preceding vowel shortens.

[26] I am grateful to Yilmaz Vural for pointing out these facts to me.

[27] Ultimately one would wish to explain why this *n* is not realised in domain-final position. This is not a property of *n*'s in general, as can be seen from the form of the ablative suffix *-dan*. To determine whether this is a lexical property of the possessive suffix or some phonological effect requires further work on Turkish.

[28] I do not consider English *keep* and *kept* to be derived from a single underlying form. See Kaye & Vergnaud (1989) for arguments.

[29] See Lowenstamm (1988) for detailed discussion of this kind of structure.

[30] Certainly, this has characterised earlier work with which I have been associated. See Kaye & Lowenstamm (1981) for one example.

[31] Transcriptions of Japanese contain word-final *n* as in *nihoN* 'Japanese'. The nasal has been termed 'syllabic' in traditional treatments. Yoshida (1989) has argued that it occupies the onset position and spreads into the final empty nucleus precisely to avoid an ECP violation.

[32] This is not the standard view. Yawelmani vowel ~ ø alternations are typically treated as epenthesis. The generalisation in (51) could then be expanded to state:

(i) In languages in which shortening takes place before a single final consonant, vowel syncope will always feed shortening and epenthesis will always bleed shortening

Or, more succinctly:

(ii) In languages in which shortening takes place before a single final consonant, this process will always be sensitive to vowel ~ ø alternations

REFERENCES

Bromberger, Sylvain & Morris Halle (1989). Why phonology is different. *LI* **20**. 51–70.

Charette, Monik (1984). The appendix in parametric phonology. *Studies in African Linguistics*. Supplement **9**. 49–53.

Charette, Monik (1988). *Some constraints on governing relations in phonology*. PhD dissertation, McGill University.

Halle, Morris & G. N. Clements (1983). *Problem book in phonology*. Cambridge, Mass.: MIT Press.

Halle, Morris & Jean-Roger Vergnaud (1978). Metrical structures in phonology. Ms, MIT.

Hayes, Bruce (1980). *A metrical theory of stress rules*. PhD dissertation, MIT.

Kaye, Jonathan (1985). On the syllable structure of certain West African languages. In D. Goyvaerts (ed.) *African linguistics: essays in memory of M. W. K. Semikenke*. Amsterdam: John Benjamins. 285–308.

Kaye, Jonathan (1990). Government in phonology: the case of Moroccan Arabic. *The Linguistic Review* **6**. 131–159.

Kaye, Jonathan (forthcoming). On the interaction of theories of lexical phonology and theories of phonological phenomena. In W. U. Dressler, H. C. Luschützky, O. E. Pfeiffer & J. R. Rennison (eds.) *Phonologica 1988*. Cambridge: Cambridge University Press.

Kaye, Jonathan & Jean Lowenstamm (1981). Syllable structure and markedness theory. In A. Belletti, L. Brandi & L. Rizzi (eds.) *Theory of markedness in Generative Grammar*. Pisa: Scuola Normale Superiore. 287–316.

Kaye, Jonathan & Jean Lowenstamm (1984). De la syllabicité. In F. Dell, D. Hirst & J.-R. Vergnaud (eds.) *Forme sonore du langage*. Paris: Hermann. 123–160.

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

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud (1989). Rektion und Konstituentrektion in der Phonologie. *Linguistische Berichte* **114**. 31–75.

Kaye, Jonathan & Jean-Roger Vergnaud (1989). Towards a theory of lexical access and lexical representation. Ms, London & Los Angeles.

Lowenstamm, Jean (1988). On the non-adjacency of geminate consonants. Paper presented at the Generative Phonology Workshop 'How different is phonology?', University of Leiden.

Lowenstamm, Jean & Jonathan Kaye (1985). Compensatory lengthening in Tiberian Hebrew. In L. Wetzels & E. Sezer (eds.) *Studies in compensatory lengthening*. Dordrecht: Foris. 97–132.

McCarthy, John (1979). *Formal problems in Semitic phonology and morphology*. PhD dissertation, MIT.

Segundo, Sylvia (1989). *Stress and related phenomena in Brazilian (Natal) Portuguese*. PhD dissertation, SOAS, University of London.

Underhill, Robert (1976). *Turkish grammar*. Cambridge, Mass.: MIT Press.

Yoshida, Shohei (1990). *Some aspects of governing relations in Japanese phonology*. PhD dissertation, SOAS, University of London.

Phonology 7 (1990) 331–35
Printed in Great Britain

*A government-based analysis of the 'mora' in Japanese**

Shohei Yoshida
Australian National University

0 Introduction

The mora is defined as something of which a long syllable ((C)VV or (C)VC) consists of two and a short syllable ((C)V) consists of one (McCawley 1968). According to some linguists, languages are classified into two groups: those that are best analysed in terms of morae and those that should be analysed in terms of syllables. Japanese is frequently cited as an example of a language that belongs to the former group. However, if the ultimate goal of the field of phonology is to seek for a group of principles that make up Universal Phonology, a unit such as the mora, which is indispensable in some languages but completely irrelevant in other languages, is an unwelcome innovation.

In this paper, I will argue that the theory of syllable structure couched within Government Phonology (Kaye *et al.* 1985, 1987) predicts that 'apparent' long syllables of the types (C)VV and (C)VN (N = the mora nasal) in Japanese cannot be analysed as consisting of single syllables. I will show instead that these strings are in fact composed of two syllables and consequently the evidence for the mora in Japanese turns out to be illusory.

This paper is organised into three major sections: §1 consists of a sketch of some motivations for the mora and the 'standard' model of Japanese syllable structure within which the strings (C)VV and (C)VN are treated as long syllables. In §2, I will give an alternative account of these strings, based on the government-based theory of syllable structure. §3 gives some empirical support for theory-internal predictions.

1 Background

1.1 Motivations for the mora

The mora is recognised as a basic unit for the description of the pitch contours of Japanese phrases. Let us look first at the process of INITIAL LOWERING (cf. Haraguchi 1977). The effect of Initial Lowering is to lower