

No initial empty CV in *clusterless* languages

Nancy C. Kula
University of Leiden/LUCL

1. Introduction

From the inception of strict CV in Lowenstamm (1996) the question of how surface clusters are licensed in Government Phonology has been a contentious one. If constituent structure consists of a series of strictly alternating C and V positions, and empty positions are licensed by proper government, then all cluster types should be available to languages that license clusters. The initial empty CV, which though originally proposed to account for certain word-initial clitic alternations in Lowenstamm (1999), emerges as a possible solution to categorizing blatant distributional facts about clusters. Viewed as a morphological (boundary) marker imported from morphology to phonology, its presence accords with languages that show restrictions on cluster type in initial position, while its absence is consistent with languages that show free variation of clusters in initial position. In the midst of this, clusterless languages are assumed to pattern with the former group in requiring the highest restriction on clusters i.e. none at all. This paper argues that there is little ground for such an assumption and that in fact, clusterless languages have no initial empty CV, a fact that is shown to follow from the general ban on proper government in this language type. The paper proceeds as follows; §2 presents the motivations for an initial empty CV; §3 presents arguments for a parametric treatment of proper-government; §4 presents empirical evidence for the absence of an initial empty CV in clusterless languages; and §5 offers some concluding remarks.

2. Why have an initial empty CV?

Lowenstamm (1996) motivates the initial empty CV on mainly two grounds. The first is the existence of languages that have initial sonority increasing clusters (henceforth TR clusters) but not sonority decreasing clusters (henceforth

RT clusters). This cross-linguistic distribution differentiates languages like French and English that allow only TR clusters in initial position from those like Moroccan Arabic and Polish that in addition quite freely allow the mirror image.

The second motivation for an initial empty CV is drawn from the fact that the initial position remains an environment that still needs to be uniquely referred to by many cross-linguistic phonological phenomena that apply exclusively at the left edge of the word.

A third motivation for the initial empty CV can be drawn from Ségéral & Scheer's (2001) discussion of the *coda mirror* where they argue for a uniform structural configuration for strong positions.

I consider these three motivations in more detail below and show why the assumption of an initial empty CV renders support for them.

2.1 TR-only languages

To begin the discussion it is necessary to understand how TR clusters, whose intervening vowel can never gain phonetic interpretation, are licensed in strict CV. As opposed to vowel~zero alternation sites, which are licensed via proper government, TR clusters are said to be in an infra-segmental government relation that renders the intervening vowel inert. In Kula (2002) I have argued that such inert vowels never project their positions to the nuclear projection and can as such not be part of the licensing relations holding within a phonological domain. Infra-segmental government proceeds from right to left and is subject to a complexity condition that requires the governor to be more complex than the governed.¹ I give the effects of the three concepts relevant to the ensuing discussion as presented in Scheer (2004) in (1) with only minor modifications, but see also Szigetvari (1999) for similar ideas.

(1) Effects of Government and Licensing

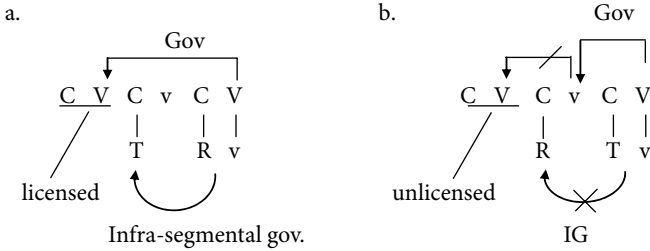
Proper Government: inhibits the segmental expression of its target

Infra-segmental government: renders an intervening vowel inert

Licensing: enhances the segmental expression of its target

Taking TR clusters to be licensed by infra-segmental government and RT clusters by proper government, the latter because they fail to satisfy the infra-segmental government conditions, languages of Indo-European descent with only TR clusters in initial position can be accounted for if an initial empty CV is assumed. Consider the illustration in (2) below. The initial empty CV is underlined in all illustrations.

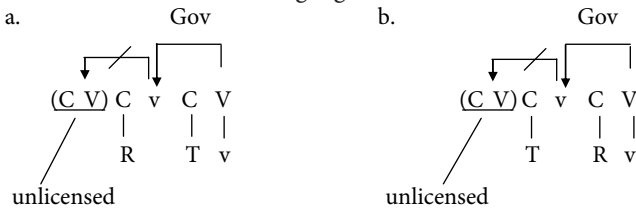
(2) TR-only languages; French, English



(2a) shows the licensing of an initial TR cluster under infra-segmental government. If an initial empty CV is present then it can be licensed to remain phonetically empty by the full vowel following such an initial TR cluster.² The distributional facts of word-initial clusters then follow from the fact that TR-only languages fail to license the initial empty CV unit whenever they are faced with an RT cluster, as illustrated in (2b). In this case infra-segmental government fails to hold and the vowel within the cluster must be licensed by proper government from the vowel following the cluster. This means that the only licenser available to the initial empty CV is the empty vowel within the RT cluster which, under standard GP assumptions, cannot license because it is uncontentful or put differently, because it is governed. In this sense TR-only languages can be categorized as having an initial empty CV.

Languages on the opposite end of the scale, on the other hand, which allow any cluster type in initial position seem to provide evidence for the absence of the initial empty CV. Consider (3) below, where both TR and RT clusters are licensed in identical fashion (by proper government) to capture the lack of co-occurrence restrictions and that in this case the initial empty CV always fails to be licensed.

(3) Both TR and RT cluster languages: Czech, Moroccan Arabic, Polish



Both (3a) and (3b) show that the initial empty CV cannot be licensed because both clusters are licensed by proper government, which renders the vowel internal to the clusters unable to license the initial empty CV. Given these facts, the initial empty CV is regarded as subject to parametric variation, present in TR-only languages but absent in languages with no restrictions on cluster type

in initial position. In this sense endorsing an initial empty CV captures an otherwise puzzling distributional fact of word-initial clusters.

2.2 Initial empty CV as a phonological alternation site

If the initial empty CV can be licensed and thereby become part of the phonological representation, we expect that, at least in some languages, the initial empty CV will act as a site of phonological alternation. Lowenstamm (1996) argues for precisely this point, showing that cliticization is enhanced by the initial empty CV. A compelling example comes from a well-documented phenomenon in Tiberian Hebrew involving the alternation between long vowels and gemination in prefixation of the singular definite article *ha-*. Consider the alternations triggered by this prefix in the minimal data in (4).

- (4) Tiberian Hebrew singular definite article alternations
- | | | | | | |
|----|------|-------|-----------|---------------------|---------------|
| a. | √dgl | degel | ‘flag’ | ha-ddegel | ‘the flag’ |
| b. | √klb | keleb | ‘dog’ | ha-kkeleb | ‘the dog’ |
| c. | √9rb | 9ereb | ‘evening’ | ha-9ereb/*ha-99ereb | ‘the evening’ |

(4a–b) show gemination of the root-initial consonant resulting from prefixation of the singular definite article, while (4c) whose root-initial guttural does not geminate shows vowel lengthening of the prefix vowel instead. As Lowenstamm (1996) shows, this distribution follows neatly if an initial empty CV is assumed as illustrated in (5a) for gemination and (5b) for vowel lengthening.

- (5) a.
The diagram shows the morphological structure of the word 'hadel'. The root 'had' (C V C V) is the base. A suffix '-el' (C V C) is attached to the root. The root is marked as 'Gov' (Government). The root is spelled out as 'h a d e' and the suffix as 'g e l'. A dashed arrow points from the 'd' in the root to the 'e' in the suffix, indicating a phonological process (deletion of the vowel in the root).
- b.
The diagram shows the morphological structure of the word 'herb'. The root 'her' (C V C V) is the base. A suffix '-b' (C) is attached to the root. The root is marked as 'Gov' (Government). The root is spelled out as 'h a r e' and the suffix as 'b'. A dashed arrow points from the 'r' in the root to the 'b' in the suffix, indicating a phonological process (deletion of the vowel in the root).

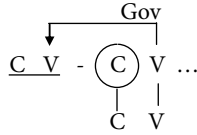
In both (5a) and (5b) the initial empty CV is licensed via proper government by the initial realized vowel of the stem, and, being phonologically licensed, can be the target of either gemination or vowel lengthening. That these are the only two logical possibilities follows directly from the representation.³

2.3 The initial position as a coda mirror/strong position

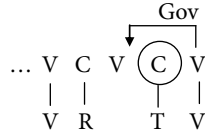
Ségéral & Scheer (2001) argue that processes that require the disjunctive context {C, #} __, i.e. processes affecting initial and post-coda positions, can uniformly be characterised as occurring in the coda mirror. Under the assumption of an initial empty CV, Scheer (2004:132) argues that their parallel phonological behaviour follows from their uniform representation as positions occurring

after a governed empty nucleus. In this sense, coda mirror (strong) positions are positions occurring in the contexts shown in (6) below.

(6) a. word-initial consonant



b. post-coda consonant



The strong C positions are presented in a circle in (6a–b). With this representation the observation that word-initial onsets and post-coda onsets behave in identical fashion gets a uniform formal representation: both occur after a governed position. For the initial position (6a), this is achieved by assuming an initial empty CV.

This falls within a wider context of formalism where weak positions are characterised as being ungoverned and unlicensed as opposed to strong positions that are ungoverned but licensed. Recall from (1) that being governed has the negative effect of inhibiting segmental expression, while being licensed has the positive effect of enhancing segmental expression. Understood in this way, ungoverned (and licensed) positions are strong while governed (and unlicensed) positions are weak. By virtue of having an initial empty CV, an initial position can escape being governed and therefore be a strong position.⁴

Thus the initial empty CV unit is motivated by the distribution of initial TR and RT clusters where languages can be distinguished with respect to whether they allow only TR clusters or any cluster type. Treating the former language type as having an initial empty CV and the latter as lacking it formally captures these distributional facts. The initial empty CV is also seen as a target of phonological processes, which is crucial in explaining the patterning of phonological alternations such as the gemination and vowel lengthening in Tiberian Hebrew. Finally, we have seen that by the assumption of an initial empty CV strong positions in word-initial and internal position can be uniformly characterised as ungoverned but licensed positions. Let us now consider the status of the initial empty CV in languages with no clusters, here dubbed *clusterless* languages.

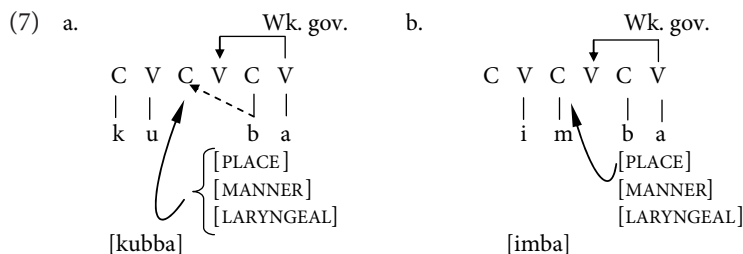
3. The status of an initial empty CV in clusterless languages

Clusterless languages are to be understood as languages that do not have true clusters in initial or any other word position. True clusters where the onsets

concerned do not share any features will be differentiated from geminates and partial geminates where either all or some features are shared between two adjacent onsets. Thus, clusterless languages such as Bantu languages, Japanese or Marshallese may have nasal-consonant clusters or geminates.

A difference between clusterless languages and true cluster languages that I would like to draw on is that clusterless languages do not license empty positions unless the flanking C positions share some features.⁵ The proposal of this paper is thus to view proper government as parametric, where proper government is to be equated to what I will term *strong* government as opposed to *weak* government. Strong government sanctions a sequence of non-dependent onsets (TR and RT clusters) while weak government applies in cases involving some dependency in terms of feature sharing (partial and total geminates). If proper government is parametric and does not function in clusterless languages, then these languages do not license empty positions by proper government and therefore cannot license the initial empty CV that requires proper government.

Alternatively, we could view the feature sharing configuration in (pseudo-) geminates as licensing the intervening empty nucleus, in which case partial and total geminates would not be licensed by proper government at all. Such a feature sharing relation could be viewed as forming a closed domain between the sharing constituents, similar to that assumed in infra-segmental government. A drawback with the latter view is that such feature/element sharing is expressed on the melodic tier and therefore does not interact with the skeletal tier. I will as such, at least for the present, remain committed to assuming strong and weak proper government. Weak government is illustrated in (7) for total and partial geminates.



Since pseudo-geminates are not allowed in initial position in many Bantu languages, it will be necessary to make further parametric assumptions on the ability of the initial nucleus which functions as the head of the domain (see Kula 2002 for arguments) to be empty. TR-only languages and the so called *anything goes* languages (exhibiting both TR and RT clusters in initial position)

can be said to have strong government manifested in surface true clusters; the RT type only in word-internal position for TR-only languages.

Under a parametric view of proper government (here strong government), if strong government is not active in a language, the initial CV cannot be licensed as the means by which it is licensed (proper government) is parametrically not available. The initial empty CV will in these cases never make the transition from morphology to phonology and is therefore not present. This is the situation that I argue holds in clusterless languages and is manifested by the lack of true clusters. It follows, then, that all other alternations that are dependent on proper government, vowel-zero alternations, for example, will also not be present in these languages. Thus while on the surface clusterless languages can be seen to have an initial realised vowel that can act as a proper-governor for the initial empty CV, and thereby pattern with TR-only languages, it fails to because strong government is not licensed. Under this view, the following categorization of languages emerges.⁶

Table 1. Cross-linguistic cluster distribution

	strong Gov	initial true clusters	internal true clusters	initial CV	Example language
Only TR	✓	✓	✓	✓	English, French
TR and RT	✓	✓	✓	✗	Czech, Polish
no TR or RT	✗	✗	✗	✗	Bantu, Japanese

From Table 1 we see that not having strong government puts a ban on all true clusters and on the initial empty CV, while having strong government can be accompanied by either licensing or not licensing of the initial empty CV. In the former case, as opposed to the latter, a restriction on initial cluster types results.

In the next section I will consider some empirical evidence from Bantu languages that supports the distributional patterns for clusterless languages as illustrated in Table 1.

4. Negative predictions of an initial empty CV

Let us consider phenomena that are not only widespread but also commonplace in Bantu, in particular; gliding accompanied by compensatory lengthening, vowel and consonant elision and the effects of prefix deletion. It will be shown that in no case can an initial empty CV be sustainable.

4.1 Gliding and compensatory lengthening

Lumasaaba (Brown 1972, Eastern Uganda), like many Bantu languages, exhibits compensatory lengthening when gliding results from two vowels coming into contact just in case the initial vowel is high. In (8) this is illustrated for nominal prefixation where *mu-* and *mi-* (here underlined) are prefixed to a vowel-initial stem.

- (8) a. u-mu-ana → umwa:na 'child'
 b. gi-mi-enja → gimye:na 'song'

Vowel hiatus being illicit in this language, the initial high vowel in the created vowel sequence glides and the second compensatorily lengthens. I give the graphic representation of this effect both when an initial CV is assumed (9b) and when it is not (9a). (As before the initial empty CV is underlined).

- (9) a.
- b.

Whether one assumes that the /u/ of the prefix partially remains in the V position or not is not at issue here, but a configuration where it is associated with the prefix onset and by which it becomes a secondary articulation on the onset is at least required to express gliding. In this sense (9a) shows compensatory lengthening into the vacated (or partially vacated) V position. (9b), with an initial empty CV (recall that the initial empty CV marks words of major categories), allows spreading into the V position of the licensed initial empty CV resulting in a hiatus between a short and long vowel which could be argued to be phonetically equivalent to a glide followed by a long vowel but would in this case fail to relate compensatory lengthening only to those environments where gliding takes place. In fact, under the assumption of an initial empty CV, we would expect all vowel-initial stems to have a long vowel, regardless of what the preceding prefix is; a fact that cannot be substantiated.

4.2 Vowel elision

In Lumasaaba, vowel deletion in a CV prefix in environments where a following stem contains a nasal-consonant sequence results in the nasal becoming syllabic, as shown in the data in (10). The prefix whose vowel deletes is underlined.

(11) a.

b.

4.3 Consonant elision

(12) Lumasaaba

a.	in-piso	→	i:pisa	‘needle’
b.	in-fula	→	i:fula	‘rain’
c.	in-nepongu	→	inewongu	‘bag’
d.	in-beba	→	imbeba	‘rat’

As should by now be expected, the facts follow directly from the absence of an initial empty CV; the deleted C makes vowel lengthening possible in (12a–b) as shown in (13a). Contrary to this, an initial empty CV would predict gemination of the stem initial C in addition to vowel lengthening (13b).

- (13) a.
- b.

Similarly, having an initial empty CV for data such as (12c–d) wrongly predicts that under consonant elision both vowel lengthening and syllabic nasal formation will occur.

- (14)

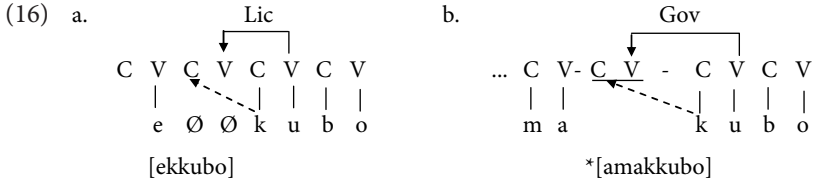
As seen in (12c–d) vowel lengthening is blocked whenever a syllabic nasal is present and while this follows from the absence of an initial empty CV, its presence would force the expression of alternations that do not occur.

4.4 Prefix deletion

We look, finally, at a prefix deletion process in Luganda that deletes the nominal class 5 prefix *-ri-*, resulting in gemination of the stem-initial onset (15a–c). This gemination can be seen as a complementary lengthening effect. The prefix is not deleted if the stem begins with a geminate (15d) and, as the plural forms of (15a–c) show, no gemination occurs with non-alternating prefixes i.e. the plural prefix *ama-* does not delete. (/l/ becomes /d/ and /y/ becomes /j/ under standard hardening processes in (15a) and (15c), respectively).

- (15) a. e-ri-lagala → eddagala ‘medicine’ ama-lagala (pl.)
 b. e-ri-kubo → ekkubo ‘road/path’ ama-kubo (pl.)
 c. e-ri-yiinja → ejjinja ‘stone’ ama-yinja (pl.)
 d. e-ri-ggwa → eriggwa ‘thorn’ ama-ggwa (pl.)

The stem-initial C gemination can straightforwardly be accounted for with no initial empty CV (16a), while assuming an initial empty CV predicts gemination even when the prefix is not deleted, as in the plural forms (16b).



Thus, based on strong empirical evidence, there is no justification for an initial empty CV in languages without true clusters because proper government is not active in these languages.⁷ As seen in the foregoing, similar conditions to those seen in Tiberian Hebrew, where gemination and vowel lengthening follow from the assumption of an initial empty CV are seen in the Bantu data, but in no case do outputs predicted by the initial empty CV actually surface.

5. Conclusion

It has been empirically shown that the initial empty CV cannot be sustained in clusterless languages and this fact has been formally related to the lack of proper government in such languages. Since the initial empty CV is licensed by proper government the absence of such a mechanism unquestionably rules out the initial empty CV. As a phonological unit, the initial empty CV captures the distributional facts of initial clusters in TR-only languages that restrict initial clusters to the sonority increasing type from those that do not and which, in addition, also allow sonority decreasing clusters. Thus, while the presence of the initial empty CV captures languages with some restriction on initial cluster types, its absence can be seen as capturing the two extreme poles of initial cluster licensing in syllable structure; either *anything goes* (both TR and RT) or *nothing goes* (no *true* clusters at all).

Notes

1. Complexity of melodic expressions is judged on number of elements.
2. Haike Jacobs points out to me that if incomplete neutralization is indeed attested in certain dialects of English where *parade* and *prayed* get identical phonetic interpretation after pre-tonic syncope, then the validity of the initial empty CV becomes questionable. In this

case, the second vowel of *parade* would properly govern the syncopated word-initial vowel, leaving no proper-governor for the initial empty CV. The matter is further compounded in cases where syncope is accompanied by liquid devoicing. I leave the discussion of this wider question to a future occasion. Readers are referred to Hooper (1978) for the relevant English facts.

3. Notice that this implies that Tiberian Hebrew falls in the category of TR-only languages that license an initial empty CV. There is compelling evidence that nouns treated as RT-initial actually betray remnants of an intervening schwa-like melody manifested in the spirantization of following stops. The details of this long and unresolved debate are beyond the scope of this paper but I thank Tobias Scheer and Delphine Seigneur for bringing these facts to my attention.
4. Notice that this entails that languages with both initial TR and RT clusters will never have word-initial strong positions. Seigneur (2003) provides positive evidence from Greek.
5. This is not to say that final empty nuclei, which are parametrically licensed, are banned in clusterless languages.
6. While initial true clusters imply internal true clusters the reverse does not hold, meaning there are languages with no initial true clusters but internal true clusters. These must have strong government (and the option of an initial empty CV). Delphine Seigneur points out to me that Saho is such a language.
7. Notice that as for languages with initial RT clusters this would predict under coda mirror assumptions that the initial position is never a strong position in clusterless languages. This is an implication that still needs empirical justification.

References

- Ashton, Ethel O., E.M.K. Mulira, E.G.M. Dawula & Archibald N. Tucker. 1954. *A Luganda Grammar*. London: Longmans, Green & Co.
- Brown, G. 1972. *Phonological rules and dialect variation: a study of the phonology of Luma-saaba*. Cambridge: Cambridge University Press.
- Hooper, Joan B. 1978. "Constraints on schwa-deletion in American English". *Recent developments in Historical Phonology* ed. by J. Fisiak, 183–207. Berlin: Mouton de Gruyter.
- Kula, Nancy C. 2002. The phonology of verbal derivation in Bemba. PhD diss., Leiden University.
- Lowenstamm, Jean. 1996. "CV as the only syllable type". *Current trends in Phonology. Models and methods* ed. by Jacques Durand & Bernard Laks, 419–441. Salford, Manchester: ESRI.
- Lowenstamm, Jean. 1999. "The beginning of the word". *Phonologica 1996* ed. by John Renison & Klaus Kühnhammer 153–166. The Hague: Holland Academic Graphics.
- Scheer, Tobias. 2004. *A lateral theory of phonology. What is CVCV and why should it be?* Berlin: Mouton de Gruyter.
- Ségéral, Philippe & Tobias Scheer. 2001. "La coda-miroir [The coda-mirror]". *Bulletin de la Société de Linguistique de Paris* 96.107–152.

- Seigneur, Delphine M. Stavrou. 2003. "Diachronic consonant lenition and exotic word-initial clusters in Greek: a unified account". *Studies in Greek Linguistics. Proceedings of the 23rd annual meeting of the department of linguistics* ed. by M. Stavrou & A. Fliatouras, 345–357. Thessaloniki: University of Thessaloniki.
- Szigetvári, Péter. 1999. VC Phonology: a theory of consonantal lenition and phonotactics. PhD diss., Eötvös Loránd University, Budapest.