EMPTY NUCLEI IN HAITIAN CREOLE*

Parth Bhatt and Emmanuel Nikiema
University of Toronto

1. Introduction and Theoretical Framework

The aim of this paper is to argue that empty nuclei are an essential part of the phonological structure of Haitian Creole (henceforth HC). Evidence for the existence of empty nuclei comes from three sources: a) the syllabification of HC consonant clusters in words such as *galri* [galʀi] "gallery" and *matlo* [matlo], "sailor"; b) the behaviour of word-final clusters and word-internal clusters in morphologically related forms as in *mèg* [mɛg] "thin" and *mègri* [mɛgʀi] "to become thin" and c) the distribution of allomorphs of the HC definite determiner.

Our analysis is presented within the framework of Government Phonology (henceforth GP, Kaye, Lowenstamm and Vergnaud, 1990 Charette, 1991 among others). GP at the syllabic level defines dependency relations within and between syllables. Furthermore, empty positions within syllabic constituents must be properly governed as defined by the Empty Category Principle (henceforth, ECP) in (1). Proper government (2) is a special government relation in which the governee is empty.

(1) *EMPTY CATEGORY PRINCIPLE*
Non final empty nuclei must be properly governed or phonetically realized

(2) *PROPER GOVERNMENT*: (Kaye, Lowenstamm & Vergnaud, 1990:219)
A properly governs B iff:
A is not empty, A is adjacent to B
There is no governing domain between A and B

We also adopt a further constraint known as Government-Licensing (3) Nikiema (1989) and Charette (1990, 1991). Governors, in order to act as such, need to be properly licensed by a phonetically realized vowel.

* We thank Pat Shaw, Carrie Dyck, and Keren Rice for very useful comments and discussion on a previous version of this paper. We cannot address all their comments here, but hope to do so in the near future. All omissions and errors are, of course, the responsibility of the authors.

(3) *GOVERNMENT-LICENSING* (Charette, 1990:242)

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

2. Consonant clusters and syllable structure in Haitian Creole

Let us now turn to a brief discussion of the questions raised by the HC data. Anestin (1987) and Cadely (1988, 1994) have convincingly argued that HC allows for branching onsets (which are always Obstruent-Liquid sequences) in word-initial position as shown in (4)[1].

(4) Word-initial obstruent-liquid clusters

flè	[flɛ]	"flower"	*grès*	[gʀɛs]	"grease"
glas	[glas]	"ice"	*tris*	[tʀis]	"sad"
plasaj	[plasaʒ]	"plant"	*krikèt*	[krikɛt]	"cricket" (insect)

The forms in (5) below contain similar clusters as those in (4) but this time in word-internal position:

(5) Word-internal obstruent-liquid clusters

tranble	[trable]	"to shake"	*enstriman*	[ɛstʀimã]	"instrument"
éklèsi	[eklɛsi]	"clearing"	*kuple*	[kuple]	"verse"

While we agree that the obstruent-liquid (henceforth OL) clusters in (4) and (5) constitute branching onsets, we contend that not all OL sequences in HC are to be analyzed as tautosyllabic. The sequences in (6) illustrate this point.

(6) Word internal obstruent liquid sequences not occuring in word-initial position

patisri	[patisʀi]	"pastry"	*sovajri*	[sovaʒʀi]	"savagery"
galri	[galʀi]	"gallery"	*matlo*	[malto]	"sailor"
bouchri	[buʃʀi]	"butcher shop"	*kozri*	[kozʀi]	"talk"
nivle	[nivle]	"to level"	*kavalri*	[kavalʀi]	"cavalry"

The OL sequences in (6) are problematic because they are not attested tautosyllabic consonants and never occur in word-initial or in word-final

[1] All the HC forms cited in this paper are taken from Valdman et al. (1996), *A learner's dictionary of Haitian Creole*.

position. The question then arises as to why these sequences show such a restricted distribution.

One solution to this problem would be to propose that these sequences of consonants do not form a branching constituent but rather a coda-onset sequence, thus explaining the absence of such sequences in word-initial position. This solution, however, leads to a violation of the Sonority Sequencing Principle (henceforth SSP) proposed by Selkirk (1982). According to this principle, the consonant in onset position cannot have a higher degree of sonority than the consonant in the preceding coda.

In order to avoid this violation, we propose that the consonant sequences in (6) are in fact separated by an intervening empty nucleus, i.e. *galri* /galØri/ and *nivle* /nivØle/. In other words, the consonants are not in fact adjacent and but rather both syllabify in onset position as shown in (7) below.[2]

(7a)

O	R	O	R	O	R	
|	|	|	|	|	|	
|	N	|	N	|	N
|	|	|	|	|	|	
x	x	x	x	x	x	
|	|	|	|	|	|	
g	a	l	Ø	R	i	[galri]

(7b)

O	R	O	R	O	R	
|	|	|	|	|	|	
|	N	|	N	|	N
|	|	|	|	|	|	
x	x	x	x	x	x	
|	|	|	|	|	|	
n	i	v	Ø	l	e	[nivle]

Within the underlying representation of forms such as *galri* (/galØRi/) or *kozri* (/kozØRi/), the final vowel properly governs the preceding empty nucleus (schwa) at the projection level. Since the empty nucleus is licensed within the structure, the nuclear position may remain without phonetic content (8).

(8)

			R ←		R	
			|		|	
O	R	O	R	O	R	
|	|	|	|	|	|	
|	N	|	N	|	N
|	|	|	|	|	|	
x	x	x	x	x	x	
|	|	|	|	|	|	
g	a	l	Ø	R	i	[galRi]

The presence of the intervening empty nucleus accounts for the absence of co-occurrence restrictions between the consonants in (6) and also accounts

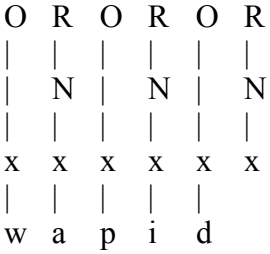
[2] Not all intervocalic consonants are separated by an intervening empty nucleus; forms such as *algèb* “algebra”, *rezilta* “result” and *èspò* “sport” have true coda-onset clusters.

for the absence of such sequences word-initially or word-finally. One should also note that a number of the words in (6) are formed by adding the nominal suffix -ri /-Ri/ to a base that ends in a consonant, i.e. *bouchri* /bu∫/+/ʀi/, *kozri* /koz/+/ʀi/. Some of these forms thus appear to be the result of affixation with base forms that end in an empty nucleus.

3. Single consonants in word-final position

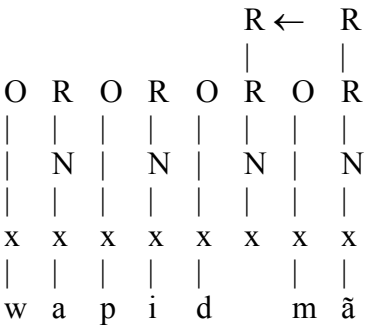
Let us first analyze lexical items with a single word-final consonant such as *jij*, *wapid*, *legal* and *piblik*. Following the analysis given in section 2, we propose that these words also contain a final empty nucleus as in (9).

(9)



When combined with a consonant-initial suffix such as *-man* [-mã], these words surface as [ʒiʒmã], [wapidmã], [legalmã] and [piblikmã]. In other words these forms do not contain an epenthetic vowel. The absence of an epenthetic vowel is explained by the fact that the empty nucleus is properly governed by the following nucleus as in (10) below.

(10)

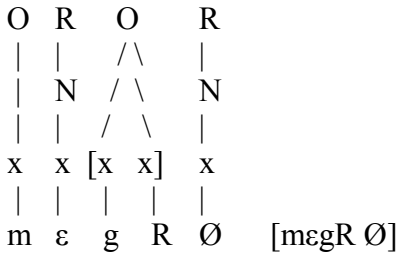


On the other hand, words such as *senp* [sɛ̃p] and *jis* [ʒis], which also appear to have a single word-final consonant, behave very differently when combined with this same suffix. The forms *senpleman* [sɛ̃**ple**mã] and *jisteman* [ʒi**ste**mã] contain a consonant cluster and an epenthetic vowel.

4. Consonant clusters in word-final position

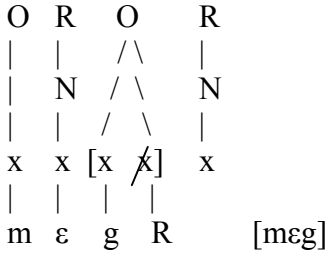
Let us begin by analyzing words such as *règ* [Rɛg] "rule", *senp* [sɛ̃p] "simple" and *mèg* [mɛg] "thin". As stated above, these forms appear to end in a single consonant. In morphologically related forms, however, such as *règle* [Rɛgle] "to rule", and *mègri* [mɛgRi] "to lose weight, become thin", two consonants occur. In order to account for this alternation, we propose that forms such as *règ* [Rɛg] "rule", *senp* [sɛ̃p] "simple" and *mèg* [mɛg] "thin" in fact contain an underlying branching onset as in (11).

(11)



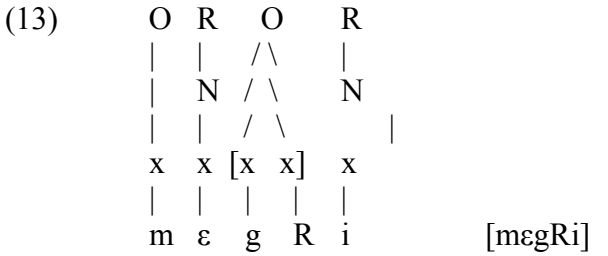
The final syllable of such words thus contains a consonant cluster followed by an empty nucleus. The explanation for the deletion of the consonant is explained by Government-Licensing (3). The word-final consonant cluster defines a governing domain that is not properly licensed. The empty nucleus can license the onset position, but the onset position cannot properly govern its complement since it is not licensed by a full vowel. The presence of an empty nucleus leads to an ill-formed syllabic structure that causes the complement position in this cluster (i.e. [R]) to be deleted. This correctly accounts for the fact that the attested surface form in HC is [mɛg] (12).

(12)



In morphologically related forms such as *mègri* [mɛgRi] "to become thin", however, the vowel of the final syllable is phonetically realized thus satisfying the condition of Government-Licensing. In this case the full vowel properly licenses the initial consonant in onset position ([g]). In turn this consonant properly governs its complement (/R/). Both consonants surface since they occur in a well-formed syllable structure (13).

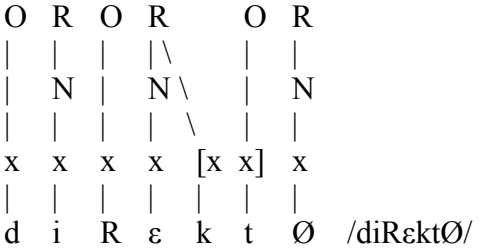
(13)



Our analysis thus correctly accounts for the deletion of the second consonant in forms where the tautosyllabic consonant cluster is followed by an empty nucleus and the retention of the tautosyllabic consonant cluster in morphologically related forms with a vowel in the following nucleus.

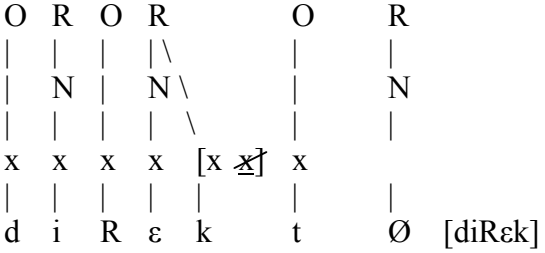
Let us now turn to words such as *jounalis*, [ʒunalis] "journalist" *datis*, [dãtis] "dentist", *korèk*, [kɔRɛk] "correct" and *dirèk*, [diRɛk] "direct", etc., where the lexical item ends in a heterosyllabic consonant cluster. Following the analysis proposed above for tautosyllabic branching clusters, we propose that such forms in fact contains an underlying coda-onset sequence as in (14) below:

(14) coda-onset sequence



As in (11), the final syllable of words such as *jounalis*, *dirèk* and *dantis* contains an empty nucleus. The empty nucleus licenses the onset position, but in this case, the onset position must govern the preceding coda. Since the onset is not properly licensed by the empty nucleus, it cannot properly govern the preceding coda and results in an ill-formed syllabic structure. It is for this reason that the onset position is deleted.

(15) coda-onset sequence

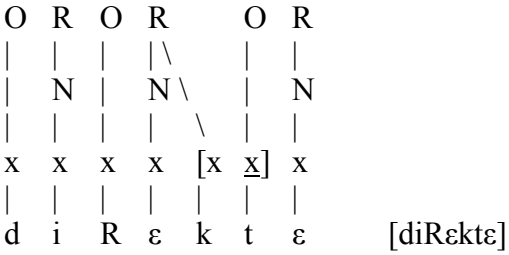


This explains why the attested forms of coda onset sequences followed by an empty nucleus contain only a single final consonant *dirèk* [diRɛk].

The analysis presented in section 4 accounts for the fact that deletion always applies to the rightmost segment within the domain. In cases such as *meg* [mɛg] the rightmost element of the domain is a complement, but in forms such as *dirèk* [diRɛk], it is a head.

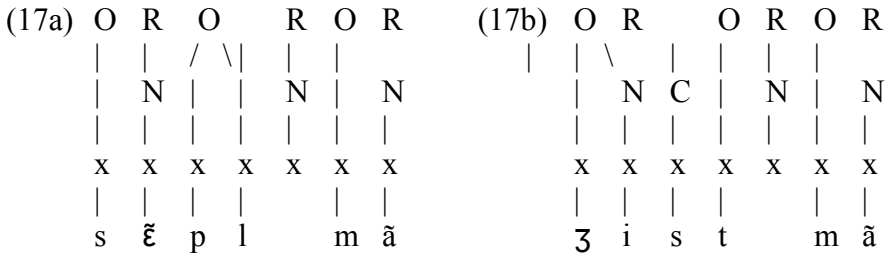
As was the case for tautosyllabic clusters, when the following nucleus contains a vowel, as in morphologically-related forms such as *dirèkte* [diRɛktɛ] "director", the onset position is licensed by the full vowel and the onset can thus properly govern the preceding coda. Since both positions are properly governed, both consonants are pronounced (16).

(16) coda-onset sequences



5. Vowel epenthesis and word final clusters

Let us now return to words such as *senpleman* and *jisteman*, where the consonant cluster is followed by a consonant-initial suffix. Following the analysis presented in section 4, we propose the following representation for these forms:



Hypothetically, these forms could have three surface realizations in HC as shown in (18a), (18b) and (18c) below. The forms in (18a) would be the result of the deletion of the second consonant. The forms in (18b) would retain both consonants and also the empty nucleus. The forms in (18c), which are the actual forms attested in HC, result from the insertion of an epenthetic vowel before the suffix. This raises the question as to why the epenthetic vowel is necessary.

(18a) *[sɛ̃pmã] (18b) *[sɛ̃plmã] (18c) [sɛ̃plemã]
*[ʒismã] *[ʒistmã] [ʒistemã]

The analysis that we have proposed up to this point predicts that the deletion of the second consonant of the root as in (18a) would result in ill-formed relations since the resulting syllabic representations (19a) clearly violate both government-licensing (the governing consonant is not properly licensed by a vowel) and (since the complement cannot be associated to a temporal position) the Empty Category Principle (locality prohibits two adjacent empty positions). The forms in (18b) are also ruled out by Government-Licensing since the empty nucleus is properly governed by the following filled nucleus but the onset consonant of the empty-headed syllable cannot govern the complement in the branching onset or the preceding consonant in coda position (19b).

(19a)

```
    O  R   O     R  O  R
    |  |  /  \   |  |  |
    |  N  |   |  N  |  N
    |  |  |   |  |  |  |
    x  x  x   x  x  x  x
    |  |  |   |     |  |
*   s  ɛ̃  p   l     m  ã
```

(19b)

```
    O  R      O  R  O  R
    |  | \    |  |  |  |
    |  N  C   |  N  |  N
    |  |  |   |  |  |  |
    x  x  x   x  x  x  x
    |  |  |   |     |  |
*   ʒ  i  s   t     m  ã
```

The only way to produce a well-formed syllabic representation with these forms is to fill the empty nucleus with an epenthetic vowel as shown in (20) below, where all dependency relations are satisfied.

(20a)

```
    O  R   O     R  O  R
    |  |  /\    |  |  |
    |  N  |   |  N  |  N
    |  |  |   |  |  |  |
    x  x  x   x  x  x  x
    |  |  |   |  |  |  |
    s  ɛ̃  p   l  e  m  ã
```

(20b)

```
    O  R      O  R  O  R
    |  | \    |  |  |  |
    |  N  C   |  N  |  N
    |  |  |   |  |  |  |
    x  x  x   x  x  x  x
    |  |  |   |  |  |  |
    ʒ  i  s   t  e  m  ã
```

Our analysis thus correctly accounts for the fact that only the forms with an epenthetic vowel ((18c), (20)) are possible surface realizations of these words. The behaviour of consonant clusters, consonant sequences and vowel epenthesis in HC clearly follow underlying structural principles that reflect dependency relations within and between syllabic constituents.

6. The forms of the HC post-posed definite determiner

Our final argument for the existence of empty nuclei in HC comes from the paradoxical distribution of the allomorphs of the definite determiner (see Nikiema 1999 and Bhatt and Nikiema 2000 for more details). The postposed HC determiner /la/ has the following six allomorphs:

(21a) **[la]** (oral consonant and oral vowel)

/ʃat + la/	[ʃatla]	"the cat"
/liv + la/	[livla]	"the book"
/bagay + la/	[bagayla]	"the thing"

(21b) **[lã]** (oral consonant and obligatory nasal vowel)

/bãk + la/	[bãklã]	"the bank"
/plãt + la/	[plãtlã]	"the plant"
/lãp + la/	[lãplã]	"the lamp"

(21c) **[na]** or **[nã]** (obligatory nasal consonant followed by an oral or nasal vowel)

/madam + la/	[madamna] or [madãmnã]	"the lady"
/wasin + la/	[wasinna] or [wasinnã]	" the root"

(21d) **[nã]** (nasal consonant and nasal vowel)

/vjann + la/	[vjãnnã]	"the meat"
/ʒanm + la/	[ʒãmnã]	"the leg"

(21e) **[a]** ("consonant deletion" followed by oral vowel)

/papa + la/	[papaa]	"the father"
/bujwa + la/	[bujwaa]	"the kettle"

(21f) **[ã]** ("consonant deletion" followed by nasal vowel)

/laʒã + la/	[laʒãã]	"the money
/dã + la/	[dãã]	"the tooth'

It appears from the above data that the phonological environment within which each of these allomorphs appears is as follows:

a) **[la]** occurs after lexical items ending in an oral vowel followed by an oral consonant (21a);
b) **[lã]** occurs after lexical items that end in a nasal vowel followed by an oral consonant (21b);

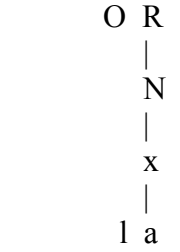
c) **[na]** can alternate with **[nã]** after words that end in a nasal consonant and which contain a vowel which subject to optional regressive nasalization (21c) (see Bhatt and Nikiema 2001) for details;
d) **[nã]** appears with lexical items that end in a nasal vowel followed by a nasal consonant (21d);
e) **[a]** appears after lexical items that end in an oral vowel (21e);
f) **[ã]** appears after lexical items that end in a nasal vowel (21f).

In other words, the consonant-initial "long" forms /la/, /na/ and /nã/ combine with words that end in a consonant, whereas the vowel-initial, empty onset "short" forms [a] and [ã] combine with words that end in a vowel. To put this paradoxical distribution another way, words that end in a consonant combine with the consonant-initial allomorph to create a sequence of two consonants, while words that end in a vowel, combine with the vowel-initial allomorph to create a sequence of two vowels.[3] The question thus arises as to why HC would create a sequence of two consonants when an empty onset allomorph is available and conversely create a sequence of two vowels when a consonant initial allomorph is available?

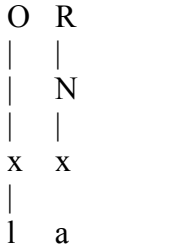
6.1 The floating consonant hypothesis

Cadely (1994) is, to the best of our knowledge, the first to represent the consonant of the determiner /la/ as a floating consonant. Cadely's representation is given in (22a), while Nikiema's (1999) representation is given in (22b). The difference between the two structures lies in the presence of a timing unit linked to the onset position.

(22a) Cadely (1994)



(22b) Nikiema (1999)



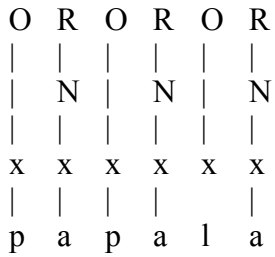
The floating consonant hypothesis is motivated by the fact that the patterns exhibited by the allomorphs of the determiner /la/ (long versus short

[3] This pattern is the exact opposite of the one observed in languages such as French, where underlying consonants are usually retained before an empty onset and dropped before a full onset in order to maintain a well-formed onset-rhyme structure (see Charette 1991).

allomorphs) can be reduced to an l/Ø alternation that is typical of floating consonants.

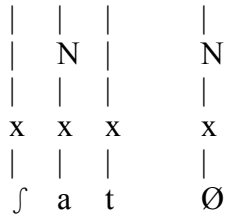
Let us begin by examining the derivation of [papaa] “the father”. In this form, the determiner follows a vowel-final stem. The empty position in the onset of the determiner is, however, properly governed by the following nucleus (vowel) and may thus remain empty. The floating consonant does not surface and a hiatus is created. Conversely, the form [papala] is ruled out because the empty position containing the floating consonant is properly governed.

(23)



Let us now analyze words such as [ʃat], “cat” which appear to have a consonant in word final position. Following the analysis presented above, we propose that all stem-final consonants in HC are not in coda position, but are in fact the onset of a syllable with an empty nucleus (24).

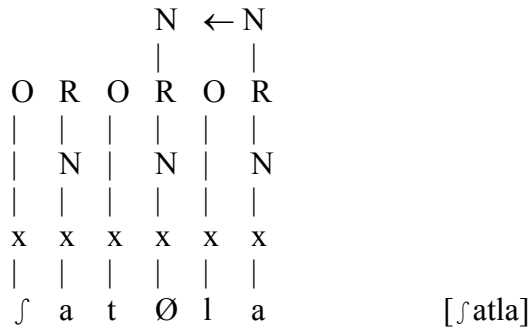
(24)



Since the empty nucleus is in word final position, it does not violate the ECP given in (2) above. When the empty nucleus of the lexical item is followed by a suffix, however, the empty nucleus must either be properly governed by a following nucleus or phonetically realized. Recall also that the underlying representation of /la/ involves an empty onset position that is subject to the ECP, i.e. the empty onset position of the determiner must also be properly governed. When the determiner /la/ follows the lexical item such as [ʃat], a sequence of two adjacent empty positions results. Because government is strictly local, our analysis predicts that only one of the two empty positions can be properly governed. One of the two empty positions will thus necessarily violate the ECP, hence ruling out the entire form. This is tantamount to saying that, if in (25) below, the final nucleus properly

governs the empty onset, the final empty nucleus of the lexical item will not be governed. The floating consonant is anchored in exactly this context to avoid an ECP violation. The floating consonant links to the empty onset (its sole available position) and leaves the government relation to occur between the two nuclei at the projection level. The non-final empty nucleus of [ʃat] is now properly governed by the vowel of [la] as illustrated in (25).

(25)



In [papaa], since the proper government relation between the empty onset and the following nucleus is linearly adjacent, it holds at the segmental level. In [ʃatla], however, the government relation is between the two nuclei at the projection level.

To sum up, the underlying representation of the determiner /la/ consists of a bipositional syllable with an empty onset position (i.e. the initial consonant is floating). In forms such as [papaa] the empty onset position is properly governed by the following vowel. In forms such as [ʃatla], suffixation of /la/ creates a sequence of two empty positions that are subject to the ECP. The floating consonant is anchored in this context to avoid an ECP violation, producing the surface form [ʃatla] and not *[ʃata].[4]

7. Conclusion

We have shown in this paper that proposing the existence of empty nuclei in HC accounts for a wide range of phenomena. Firstly, the presence of a word-internal empty nucleus accounts for the fact that consonants that appear to be adjacent as in *kozri* "talk" and *nivle* "to level" are in fact the onsets of two syllables and therefore do not violate the Sonority Sequencing Principle. Secondly, proposing a word-final empty nucleus explains why forms such as *règ* "rule" and *dirèk* "direct" are pronounced with a single

[4] When combined with the definite determiner, forms which end in a consonant cluster such as *liv* [liv] "book" are pronounced as *livla* [livla], not **livrela* *[livRela] or **livrea* *[livRea]. This indicates that the government relation is between the nuclei at the projection level.

word-final consonant, while morphologically related forms such as *règle* "to rule" and *dirèkte* "director", which contain a final vowel, are pronounced with both consonants. The presence of the word-final empty nucleus accounts for the deletion of the rightmost consonant in both types of clusters. Thirdly, the existence of a word-final empty nucleus explains why forms such as *jijman* and *wapidman* combine directly with the consonant initial affix -man and do not contain an epenthetic vowel, while *direkteman* and *senpleman* contain an epenthetic vowel. Finally, the presence of empty nuclei accounts for the paradoxical distribution of the definite determiner in HC, which creates sequences of consonants or vowels. The empty nucleus of the determiner is properly governed in vowel-final stems, but forces the anchoring of the initial consonant of the determiner to avoid a violation of the Empty Category Principle in consonant-final stems.

In short, in HC, which is traditionally thought to have little or no productive morphology, word-edge morphophonological processes are very intricate and crucially involve empty nuclei and structural dependencies within and between syllables. Our analysis has shown that apparent surface violations of syllabic structure and paradoxical morphophonological processes can be explained in a principled manner by proposing empty positions and examining underlying, abstract constituents and relations.

References

Ans, A.-M. d' 1968. *Le créole français d'Haiti*. The Hague: Mouton.

Anestin, A. 1987. *Structure syllabique de l'haïtien et nasalisation*. Master's Thesis, U.Q.A.M.

Bhatt, P. and E. Nikiema 2001. 'Le statut de la nasalité en créole de Sainte Lucie', *Revue québécoise de linguistique*, 28, pp.23-45.

Cadely, J.-R. 1988. 'Représentations syllabiques et distribution des diphtongues en créole haïtien', *Études créoles* , XI (1), pp.9-40.

Cadely, J.-R. 1994. *Aspects de la phonologie du créole haïtien.* Thèse de doctorat, U.Q.A.M..

Cadely, J.-R.1996. 'On Nasality in Haitian Creole: the Case of the Definite Determiner', paper presented at the *11th Biennial Conference, Society for Caribbean Linguistics.*

Charette, M. 1991. *Conditions on Phonological Government.* Cambridge: C. U. P.

Kaye, J., J. Lowenstamm and J.-R. Vergnaud 1990. 'Constituent Structure and Government in Phonology', *Phonology,* 7:2, pp.193-231.

Nikiema, E. 1989. 'Gouvernement propre et licenciement en phonologie: le cas du tangalé', *Langues orientales anciennes, philologie et linguistique*, 2, pp.225-251.

Nikiema, E. 1999. 'De la variation morphophonologique du déterminant /la/ dans les créoles haïtien et st-lucien', *Lingua*, 107, pp.69-93.

Selkirk, E. 1982. 'Syllables', in *The structure of phonological representations*, H. van der Hulst and N. Smith (eds.), Dordrecht: Foris, pp.337-383.

Tinelli, H. 1981. *Creole Phonology*. The Hague: New York: Mouton.
Valdman, A. 1978. *Le créole: statut et origine*. Paris: Klincksieck.
Valdman, A., C. Pooser and R. Jean-Baptiste 1996. *A learner's dictionary of Haitian Creole.* Bloomington: Creole Institute.