How complex is Creole Inflectional Morphology?

The case of Mauritian

Olivier Bonami & Fabiola Henri
Université Paris-Sorbonne, LLF, Institut Universitaire de France
olivier.bonami@paris-sorbonne.fr
Laboratoire de Linguistique Formelle (UMR7110 – U. Paris Diderot & CNRS)
henrifabiola@gmail.com

14th International Morphology Meeting, Budapest, May 2010

1 Introduction

1.1 The issue: creole complexity

- It is now well established that Creole languages can have morphology, but it is expected to be 'simple', 'easy', 'unmarked', 'natural', or 'canonical' (see Plag, 2006, and references therein).
  - ☞ In the domain of inflectional morphology, we expect affixal morphology to express clearcut morphosyntactic features with no inflectional classes and little irregularity.
- We address this issue in detail in the case of Mauritian, a French-based Creole.
- We show that although Mauritian inflection is simple with respect to paradigm size, it is complex in other respects. In particular,
  - ☞ the interface with syntax and semantics is not simple and
  - ☞ the inflectional system as a whole is not easier to use.

1.2 Mauritian morphology

- Most of the language's vocabulary has been inherited from French with a few phonological adaptations.

French↦Mauritian	example		trans.
ʃ↦s	detaʃe	↦detase	'detach'
ʒ↦z	mɑ̃ʒe	↦mɑ̃ze	'eat'
ʁ↦ɚ/__[$_\sigma$	paʁti	↦paɚti	'leave'
y↦i	fyme	↦fime	'smoke'
ə↦e/#C__	ʁədone	↦ʁedone	'give again'
ɛ↦e	fɛʁ	↦feɚ	'do'
ɔ↦o	sɔʁti	↦soɚti	'go out'

- It has however not inherited from the French inflectional system. In fact, Mauritian show no inflection w.r.t. tense, mode and aspect or to number and person.

(1) a. Mo/to/li/nou/zot manz kari.
1SG/2SG/3SG/1PL/2/3PL eat.SF curry
'I/you/he/she/they eat(s) curry.'

b. Mari/Zan manz kari.
Mary eat.SF curry
'Mary eats curry.'

(2) a. Mo ti manz kari.
1SG PST eat.SF curry
'I ate curry.'

b. Mo pou manz kari.
1SG PST eat.SF curry
'I will eat curry.'

- Still, most Mauritian verbs have two forms: the long form (LF) and the short form (SF).
  - ☞ The LF almost always derives from the Fr. infinitive or past participle (Veenstra, 2004)
  - ☞ The SF often resembles a Fr. present singular

LF	aʁete	amɑ̃de	apaʁet	atan	bɑ̃de	ban	bʁije	bʁije
SF	aʁet	amɑ̃d	apaʁet	atan	ban	ban	bʁij	bʁije
INF.	**aʁete**	**amɑ̃de**	**apaʁɛtʁ**	**atɑ̃dʁ**	**bɑ̃de**	*English*	**bʁije**	**bʁuje**
PPE.	**aʁete**	**amɑ̃de**	apaʁy	atɑ̃dy	**bɑ̃de**	*English*	**bʁije**	**bʁuje**
TRANS.	'stop'	'amend'	'appear'	'wait'	'bandage'	'ban'	'glow'	'mix'

LF	demɑ̃de	egziste	feɚ	feʁe	fini	fʁiz	fʁize	kɔ̃siste	ʁesikle
SF	deman	egzis	feɚ	feɚ	fini	fʁiz	fʁiz	kɔ̃siste	ʁesikle
INF.	**dəmɑ̃de**	**ɛgziste**	**fɛʁ**	**feʁe**	finir	*English*	**fʁize**	**kɔ̃siste**	**ʁəsikle**
PPE.	**dəmɑ̃de**	**ɛgziste**	fɛ	**feʁe**	**fini**	*English*	**fʁize**	**kɔ̃siste**	**ʁəsikle**
TRANS.	'ask'	'exist'	'do'	'shoe'	'finish'	'freeze'	'curl'	'consist'	'recycle'

LF	kuveɚ	mɔ̃tʁe	mine	paɚti	pase	paste	poɚte	sɑ̃ti
SF	kuveɚ	mɔ̃tʁe	min	paɚti	pas	*pas*	poɚt	sɑ̃ti
INF	kuvʁir	**mɔ̃tʁe**	**mine**	paʁtir	*English*	*Mauritian*	**pɔʁte**	sɑ̃tir
PPE.	**kuvɛʁ**	**mɔ̃tʁe**	**mine**	**paʁti**	*English*	*Mauritian*	pɔʁte	**sɑ̃ti**
TRANS.	'cover'	'show'	'undermine'	'leave'	'pass'	'filter'	'carry'	'feel'

LF	soɚti	vɑ̃de	vini
SF	soɚt	van	vin
INF.	sɔʁtiʁ	**vɑ̃dʁ**	vəniʁ
PPE.	**sɔʁti**	vɑ̃dy	**vənu**
TRANS.	'go out'	'sell'	'come'

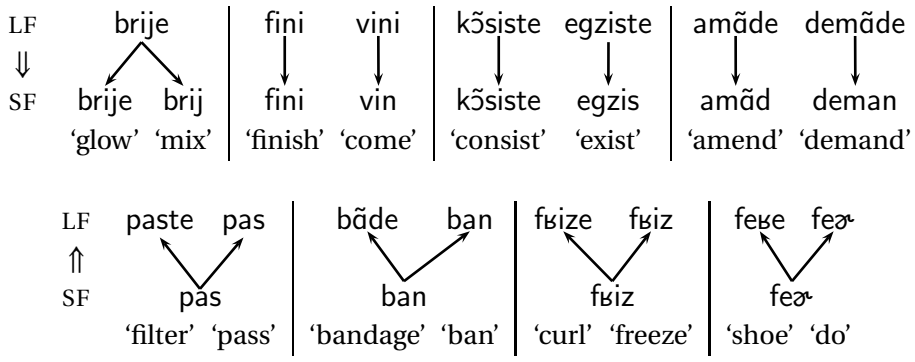
1.3 The data set

- The data is based on Carpooran (2009), the first monolingual dictionary of Mauritian with standardized orthography, phonological and etymological information.
- 2079 phonologically distinct verbal entries.
- 30% have syncretic LF and SF.
- 1850 inherited from French verbs, 34 from other categories, the rest not inherited from French

2 Why morphology?

2.1 The alternation is not phonologically predictable

- Neither form is uniformly predictable from the other:



- **Deriving the SF form the LF:**
  - Verbs with a LF in -e tend to drop the final vowel when it is preceded by a single consonant.
  - -e never drops after a branching onset.
  - Both situations are found when the verb penultimate syllable has a nonempty coda (kɔ̃siste vs. egziste) or when the single consonant is a glide (brije'mix' vs. brije 'glow').
  - Almost all verbs with a LF in -i are syncretic, but there are two exceptions (soɔ̯ti and vini), which are not phonologically distinguishable from syncretic verbs (resp. paɔ̯ti and fini).
  - Only verbs with a final consonant in the LF are uniformly syncretic.

- **Deriving the LF form the SF:**
  - verbs with a vowel-final SF are always syncretic.
  - verbs with a consonant-final SF may have a syncretic LF, a LF in -e or a LF in -i: brize vs friz, arete vs aparet, mine vs vini, poɔ̯te vs soɔ̯ti

2.2 The distribution is morphomic

The contexts in which the two forms appears does not form a natural class (Henri and Abeillé, 2008; Henri, 2010) and in lexeme formation processes both forms are used in a way that does not reflect any morphosyntactic property (Henri, 2010).

Syntactic import

- The SF is triggered by nonclausal complements (3).

(3) a. Mo ti manz/*manze kari.
1SG PST eat.SF/LF curry
'I ate curry.'

b. Sa stati la dat/*date depi lepok lager.
DEM statue date.SF/LF from period war
'This statue dates back from the war period.'

- It also appears with predicative APs (4a) and locative goals (4b).
- Verbs with a clausal complement take a SF only if another nonclausal complement precedes it (4c).

(4) a. Nou res/*reste malad.
1PL stay.SF/LF sick
'We are still sick.'

b. Li pe mars lor disab.
3SG.M PROG walk.SF on sand
'He is walking towards the sand.'

c. Mari inn demann/*demande [ar tou dimounn] [kiler la].
Mary PERF ask.SF/LF with all people what_time DEF
'Mari asked everyone what time it was.'

- Finally note that the postverbal argument of unaccusative verbs counts as a complement (5)

(5) Inn ariv/*arive enn aksidan.
prf arrive.SF/LF INDF accident
'There has been an accident.'

- Conversely, the LF appears when the verb has no complement, (6a), the complement is extracted (6b), or it is clausal (6c).

(6) a. Mo ti manze/*manz.
1SG PST eat.LF/SF
'I ate.'

b. Tibaba ki mo mama ti veye/*vey toule zour.
little_baby COMP POSS mother PST look_after.LF/SF every day
'It's little babies that my mother looked after every day.'

c. Mari inn demande/*demann [kiler la] [ar tou dimounn] .
Mary PERF ask.SF/LF what_time DEF with all people
'Mari asked everyone what time it was.'

- Adjuncts also trigger the LF.

(7) Li pe marse lor disab.
3SG.M PROG walk.LF on sand
'He is walking on the sand.'

- The alternation is **not** phonologically conditioned: a complement that is not adjacent to the verb still triggers the SF.

(8) Nou res/*reste toultan malad.
1PL stay.SF/LF always sick
'Lit. We remain always sick.'

Discursive import

- Interestingly, the LF may appear with a nonclausal complement under certain discursive conditions, precisely in counter-oriented moves (deferments, counter-implicative and counter-propositional moves).

  ☞ In such contexts, the LF is analyzed as an exponent of Verum Focus (Henri et al., 2008; Henri, 2010).

(9) Mo ti krwar Mari pa MANZE/*MANZ kari poul!
1SG PST think Mary NEG eat.LF/SF curry chicken
'I thought Mary DIDN'T eat chicken curry!'

Morphological use

- The two forms are used in "attenuative" reduplication which is a derivational process creating new verbal lexemes (Henri, 2010).

  ☞ These reduplicated form's SF is the concatenation of two copies of the base's SF, whereas the LF is the concatenation of the base's SF with the base's LF.

LF	SF	gloss	red. LF	red. SF	trans.
sɑ̃te	sɑ̃t	'sing'	sɑ̃tsɑ̃te	sɑ̃tsɑ̃t	'hum'
reste	res	'stay'	ʁesʁeste	ʁesʁes	'stay occasionally'
soɔ̰ti	soɔ̰t	'get out'	soɔ̰tsoɔ̰ti	soɔ̰tsoɔ̰t	'get out occasionally'
balje	balje	'sweep'	baljebalje	baljebalje	'sweep carelessly'

Examples of attenuative reduplication

- Attenuative reduplication contrasts with intensive reduplication, which is a syntactic rather than a lexical process, and where both the base and the reduplicant are always exact copies.

(10) a. Mo ti manze, manze, manze.
1SG PST eat.LF eat.LF eat.LF
'I ate, ate, ate.'

b. Zan nek sant sega, sant sega mem enn lazourne.
John only sing.SF sega sing.SF sega still day
'John keeps singing the sega, singing the sega all day long.'

Conclusion

- Mauritian verbs **do** have (tiny) inflectional paradigms
- The distribution of the forms is morphomic (Aronoff, 1994):
  - Both forms are used both in inflection and lexeme formation.
  - Neither form realizes a constant set of features.

3 How complex morphology?

3.1 Method

We want to address two issues:

- Whether one of the two forms of the Mauritian verb should be conceived as the 'base form'.
- Whether Mauritian inflection is simpler than French inflection is other respects besides paradigm size.

In both cases we need a measure of the complexity of the morphophonological derivation relation between two cells in the paradigm:

- Within Mauritian, if the relation between the two cells is 'simpler' in one direction than in the other, the more predictive cell should be considered the base form (Albright 2003, Blevins 2006).
- To compare French and Mauritian, we can compare how hard it is to predict comparable cells from comparable cells in each language's paradigm.

Among the many conceivable strategies, we retained 4 ways of assessing the complexity of the (oriented) relation between two cells.

3.1.1 Number of undecidables

We call *undecidable* a situation where two distinct lexemes have identical forms in the input cell but different forms in the output cell. Examples are given below.

language	relation	input	output #1	trans.	output #2	trans.
Mauritian	LF⇒SF	bʁije	bʁij	'glow'	bʁije	'scramble'
Mauritian	SF⇒LF	ban	bɑ̃de	'bandage'	ban	'ban'
French	PST.PCPL⇒PRS.SG	ply	plø	'rain'	plɛ	'please'

Undecidables are a problem for any strategy for describing the relation between two forms: any strategy will make the wrong prediction for at least one of the two forms. Thus the more undecidables there are, the more complex the relation is.

3.1.2 Number of distinct morphophonological pattern

By a morphophonological pattern we mean a modification of the input form that can be described as the structural change of a phonological rule in SPE format (Chomsky & Halle, 1968). Thus different rules active in an inflectional system may share the same morphophonological pattern. As an example here are the patterns for the Mauritian LF⇒SF relation:[1]

(11) a. V⟶ ∅
b. X⟶X
c. ʁe⟶ɚ

[1]Notice that we treat [ɚ] and [ʁ] as realizing distinct phonemes, despite the fact that they are in complementary distribution and both derive from French /r/. Contrary to what one might expect, this makes for a tighter description of Mauritian , because [ɚ] more often falls in a natural class with vowels than with consonants. However a consequence of this move is that the number of morphophonological changes is a bit higher than expected.

d. Ce⟶ ∅
e. V[+nas] C[−cont,+voice,−nas]e⟶ V[−nas]O[+nas]
f. ble⟶m
g. ɔ̃jate⟶ʁjat

The number of morphophonological patterns is sensitive to both the number of inflectional patterns and the diversity of suppletive relations.

3.1.3 Reliability of the best ECNR

By an Exhaustive Categorical Nonlexical Rule (ECNR) we mean a set of morphophonological rules such that

- All rules are mutually exclusive: for a given phonological input, at most one rule may apply.
- The rule system is exhaustive: for all phonotactically well-formed input, some rule may apply.
- No rule is *lexical*, in the sense that it applies to a single input form in the data set.

There are many different ECNRs for a given set of pairs of forms. We can compare them in terms of (raw) reliability:

(12) The (raw) reliability of a morphophonological rule is the ratio between the number of succesful applications (hits) and the number of possible applications (scope).

$$\text{raw reliability} = \frac{\text{hits}}{\text{scope}}$$

Since ECNRs are intended to be exhaustive, the scope is the total number of pairs of forms in the data set. For example, the following ECNR for the Mauritian LF⇒SF relation has a reliability of 98.16%.

(13) If the LF is polysyllabic and ends in VCe, drop the e; otherwise do nothing.

The *best* ECNR for a given relation is the ECNR with the highest reliability. For example, here is a specification of the best ECNR for the Mauritian LF⇒SF relation (reliability: 99.09%).[2]

(14) a. e⟶ ∅/V[−nas,−rhot]C_
b. e⟶ ∅/V[+nas,−low]C_
c. e⟶e/CV_
d. e⟶e/CC_
e. e⟶e/#C_
f. ãde⟶an
g. e⟶ ∅/ãC[−cor]_
h. e⟶ ∅/ɔ̃C[−voc]_
i. e⟶e/ɔ̃C[+voc]_
j. X⟶X, where X ≠ e

The best ECNR can be conceived as a statement of regular morphology under a maximally inclusive view of regularity: it is the best one can put into a rule system short of including individual lexical entries.

[2] For now we compute the best ECNR by hand, which is a long and cumbersome process on data sets with over 2000 pairs. Thus it might be that we did not succesfully identify the absolute best rule.

3.1.4 Reliability of the MGL

- The *Minimal Generalization Learner* (Albright, 2002) is a stochastic model of inflection based on SPE-style morphophonological rules.
- The MGL relates two paradigm cells:
  - First, rules are learned from a training corpus in a step by step process
  - Each rule is assigned a reliability measure taking into account the scope of the rule.
  - The system predicts multiple candidate outputs for a nonce or existing base, and ranks the outputs on the basis of reliability.

The learning algorithm

- The training corpus consists of pairs of forms: ⟨base, target⟩
- For each pair, the MGL starts by determining the minimal rule relating the two inflected forms. The rule consists of a structural change and a context in the SPE tradition
- Example
  - pair to be learned: ⟨lave, lav⟩
  - structural change: e→ ∅
  - context: #lav_#
  - minimal rule: e→ ∅/#lav_#
- This new rule is compared to previously learned rules, in order to make minimal generalizations.
  - ☞ If two rules have the same structural change and both their contexts are subsumed by a more general context of the form

    $$\#\ (X)\ [\text{feat}^+]\ \text{seg}^*\ \_\ \text{seg}^*\ [\text{feat}^+]\ (Y)\ \#$$

    infer a new rule with that more general context.
  - Example: from e→ ∅/#lav_# and e→ ∅/#gav_#, infer

    $$\text{e} \rightarrow \emptyset / \#[+\text{cons}, -\text{vow}, +\text{voice}]\text{av}\_\#$$
- The output of the learning algorithm is a set of rules that is always larger than the set of input pairs, potentially much larger.

Making predictions

- Each rule the system learned is assigned an adjusted reliability, using confidence limit statistics. Specifically:

$$\text{adjusted reliability} = \text{raw reliability} - z_{0.125} \times \sqrt{\frac{\text{raw reliability} \times (1 - \text{raw reliability})}{\text{scope}}}$$

☞ At the same raw reliability, the adjusted reliability increases with the scope of the rule.

- For any candidate input, the MGL determines the list of rules matching that input. The predicted output is the output of the rule with the highest adjusted reliability.

Our use of the MGL

- The MGL has been argued to mirror satisfactorily some aspects of human performance in inflection of nonce forms (Albright, 2003; Albright & Hayes, 2003).
- We assume that it is a good enough approximation of a human learner to give us some indication as to whether some inflectional relation is easy or hard to learn and use.
- Thus we compare the reliability of the MGL on different data sets and compare the results. Specifically:
  - In each case we trained the MGL on 2079 pairs.
  - We then determined the reliability of the MGL's prediction on those same 2079 inputs.

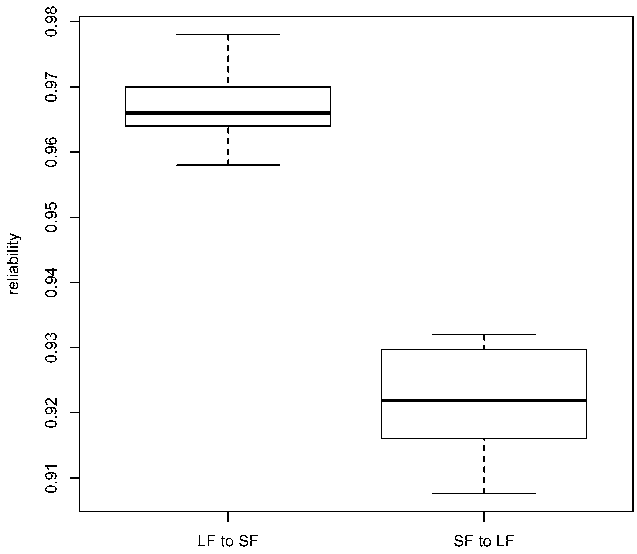
  ☞ One might think that testing the system's reliability on the same forms it was trained on does not provide a realistic estimation of its reliability when applied to nonce forms. In fact this is not the case for the data at hand.
    - The reliability of the MGL on the Mauritian LF⇒SF trained on 2079 pairs and tested on the same 2079 inputs, is 96.82%.
    - We also checked the reliability of the MGL when trained on a random set of 1300 pairs, and tested on a random set of 500 distinct forms.
    - Over 25 iterations, the mean reliability is 96.70%, standard deviation 0.65%

4 Determining the base form

	LF⇒SF	SF⇒LF
# of undecidables	1	13
# of patterns	7	10
Reliability of best ECNR	99.09%	90.43%
Reliability of the MGL	96.82%	93.18%

In each case we observe that the LF⇒SF relation is more reliable than the SF⇒LF relation

- There is a sharp contrast between the number of undecidables in the two directions. This strongly indicates that predicting the LF from the SF is the safer route.
- The numbers of patterns are not different enough that much can be concluded.
- There is a sharp contrast between reliability of the best ECNRs. This is mainly because from a consonant-final SF, there is no reliable way of deciding whether the LF will add final e or not.
- The contrast between the reliability of the MGL in both directions is not that sharp. This is because the MGL learns many details about individual pairs that the ECNR is designed to ignore.
- However we verified that this difference is statistically significant. To do this we chose a random set of 1300 lexemes in the data set, trained the MGL over the pairs ⟨LF, SF⟩ from this set, and then tested it on 500 LFs from a separate random set. We then used the inverse ⟨SF, LF⟩ over the same set of 1300 lexemes, and tested the result over the matching 500 SFs. The results after 10 iterations are summarized in the boxplots below. A paired t-test confirmed that the difference is significant, $p < 10^{-6}$.



Conclusion: It is a matter of dispute whether paradigms should be assumed to contain a fixed base cell (Albright, 2002a; Blevins, 2006; Bonami & Boyé, 2007). **If** one wants to argue that there is an oriented relationship between the two forms, the LF should be taken as the base form.

5 Comparison with French

- We now want to compare the complexity of the Maurician inflection system with relevant subparts of the French inflection system.
- Since the LF usually is inherited from a French infinitive or past participle and the SF resembles a present singular, the most relevant comparisons are:
  - Mauritian LF⇒SF with French INF⇒PRS.SG
  - Mauritian LF⇒SF with French PST.PCPL⇒PRS.SG
- Since we have a data set of 2079 verbs in Mauritian, we constructed a parallel data set of forms from the 2079 most frequent French verbs[3].

	LF⇒SF	INF⇒PRS	PST.PCPL⇒PRS
# of undecidables	1	0	1
# of patterns	7	20	26
Reliability of best ECNR	99.09%	96.92%	96.49%
Reliability of the MGL	96.82%	96,27%	95,19%

Overall, the results show very little difference between the Mauritian and French data.

- The number of undecidables is too low to be meaningful
- There are more patterns in French.
  - Reflects a higher number of small classes of lexemes with an idiosyncratic pattern.
  - Many of these lexemes have not been inherited by Mauritian

[3]Lemma frequencies in written French from version 3 of the *Lexique* database (New *et al.*, 2001). We corrected the phonological transcriptions by hand and removed entries which are not distinguishable on the basis of the three paradigm cells at hand.

– Those that have are regularized, e.g.

French			Mauritian		
INF	PST.PCPL	PRS.SG	LF	SF	trans.
ale	ale	va	ale	al	‘go’
vəniʁ	vəny	vjɛ̃	vini	vin	‘come’
dəvwaʁ	dy	dwa	dwa	dwa	‘must’
valwaʁ	valy	vo	vo	vo	‘be worth’

☞ Although Mauritian has irregular morphology, it is **not** inherited irregularity

- The best ECNRs are a little less reliable in Mauritian than in French.
  - This small difference is due to a larger number of small clusters (often singletons) of lexemes with an idiosyncratic pattern but whose forms do not contrast phonologically from those involved in other patterns. Some examples follow.

dominant pattern	example	exceptions
əne⟶ɛn	⟨amənə, amɛn⟩ ‘take to’	⟨deʒənə, deʒœn⟩ ‘have lunch’
ije⟶i	⟨tʁrije, tʁi⟩ ‘sort’	⟨bʁije, bʁij⟩ ‘glow’
aiK⟶ai	⟨ɑ̃vaiʁ, ɑ̃vai⟩ ‘invade’	⟨aiʁ, ɛ⟩ ‘hate’
e⟶ ∅	⟨ʁale, ʁal⟩ ‘moan’	⟨ale, va⟩ ‘go’

  - However since the overall type frequency of these clusters is quite small, the French ECNRs are still quite reliable.

- The MGL results are barely distinguishable, in particular for LF⇒SF vs. INF⇒PRS.

  ☞ The MGL is less sensitive to small clusters of idiosyncratic lexemes, because it records lexical information.

  – Since the MGL encodes a more realistic general model of inflection, these results show convincingly that the overall system of Mauritian has the same level or reliability as the corresponding subpart of the French system.

  – Notice that for both languages the MGL misses important generalizations, but the number of misses is comparable.

    * For Mauritian LF⇒SF, the MGL misses the generalization that consonant clusters preclude the drop of e; this accounts for 28 of the 66 errors.
    * For both French relations, the MGL can not generalize from individual rules such as əve→ɛv/X_# and əte→ɛt/X_# to əCe→ɛC/X_#; this accounts respectively for 31 of the 69 errors when starting from the infinitive and 32 of the 89 errors when starting from the past participle.

6 Conclusions

Creole complexity Mauritian Creole inflection is simpler than French inflection in some respects only:

- Paradigms are smaller
- There are fewer distinct inflection patterns

- The relationship between morphosyntactic properties and exponents is just as opaque
- Predicting a paradigm cell on the basis of another is **not** simpler

☞ Presumably, Mauritian inflection is neither easier to learn nor easier to use.

Morphological theory
Because it exhibits complex morphology despite its tiny paradigms, Mauritian inflection provides an interesting line of argument in favor of an abstractive ('word-based') view of morphology (Blevins, 2006).

- Despite tiny paradigms the system relies heavily on implicational morphology
- The base form is the more complex form
- Postulating stems would complexify the description without adding any insight
- There is word-based lexeme formation

References

Albright, A. (2003). Islands of reliability for regular morphology: Evidence from italian. *Language*, 78:684–709.

Albright, A. C. (2002). *The Identification of Bases in Morphological Paradigms*. PhD thesis, University of California, Los Angeles.

Albright, A. C. and Hayes, B. P. (2003). Rules vs. analogy in english past tenses: A computational/experimental study. *Cognition*, 90:119–161.

Aronoff, M. (1994). *Morphology by itself*. MIT Press, Cambridge.

Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, 42:531–573.

Bonami, O. and Boyé, G. (2007). Remarques sur les bases de la conjugaison. In Delais-Roussarie, E. and Labrune, L., editors, *Des sons et des sens*, pages 77–90. Hermès, Paris. Ms, Université Paris 4 & Université Bordeaux 3.

Carpooran, A. (2009). *Diksioner Morisien*. Koleksion Text Kreol, Sainte Croix (Mauritius).

Corne, C. (1982). The predicate in Isle de France Creole. In Baker, P. and Corne, C., editors, *Isle de France Creole. Affinities and Origins*, pages 31–48. Karoma, Ann Arbor.

Henri, F. (2010). *A Constraint-Based Approach to Verbal Constructions in Mauritian: Morphological, Syntactic and Discourse-based Aspects*. PhD thesis, University Paris Diderot- Paris 7 and University of Mauritius.

Henri, F. and Abeillé, A. (2008). Verb form alternations in Mauritian. In Müller, S., editor, *Proceedings of the 15th Conference on HPSG*, pages 378–398, Stanford. CSLI Publications.

Henri, F., Marandin, J.-M., and Abeillé, A. (2008). Information structure coding in Mauritian: Verum Focus expressed by long forms of verbs. Paper presented at the Workshop on Predicate Focus, Verum Focus, Verb Focus.

New, B., Pallier, C., Ferrand, L., and Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: Lexique. *L'Année Psychologique*, 101:447–462.

Plag, I. (2006). Morphology in Pidgins and Creoles. In Brown, K., editor, *Encyclopedia of Language and Linguistics, 2nd Edition*, volume 8, pages 304–308.

Veenstra, T. (2004). What verbal morphology can tell us about Creole genesis: the case of French-related Creoles. In Plag, I., editor, *Phonology and Morphology of Creole Languages*, number 478 in Linguistische Arbeiten. Max Niemeyer Verlag Gmbh.

Veenstra, T. and Becker, A. (2003). The survival of inflectional morphology in French-related Creoles. *SSLA*, 25:285–306.