

# Tracing the origins of inflection in creoles

## A quantitative analysis

Olivier Bonami<sup>1</sup>   Fabiola Henri<sup>2</sup>   Ana R. Luís<sup>3</sup>

<sup>1</sup>U. Paris-Sorbonne & IUF & LLF  
olivier.bonami@paris-sorbonne.fr

<sup>2</sup>U. Lille 3 & STL & LLF  
henrifabiola@gmail.com

<sup>3</sup>Universidade de Coimbra & CELGA  
aluis@fl.uc.pt

Ninth Creolistics Workshop, Aarhus, April 2012

# Outline

Introduction

Comparing lexifier to creole  
Indo-Portuguese  
Mauritian

Towards explaining the differences

The origin of Mauritian verb forms

Conclusions

References

# Introduction

- ▶ Much previous work on morphology in creoles focuses on
  - ▶ What morphology (if any) do creoles have
  - ▶ Do creoles tend to have 'less morphology' than their lexifier, and if so, why
- ▶ Different question: given that some creoles have nontrivial inflectional morphology, why do they have the type of morphology they have?
- ▶ We compare the conjugation system of two Indo-Portuguese creoles (Daman and Korlai) with that of Mauritian.
- ▶ We argue that
  1. Statistically prevalent features of the lexifier system shape the creole system
  2. This is partly independent of the actual forms the creole inherits
- ▶ Heavy use of quantitative data on the lexifiers and, where available, on the creoles

# Sources

- ▶ Linguistic descriptions
  - Daman & Korlai Clements (1996); Clements and Koontz-Garboden (2002)
  - Mauritian Henri (2010)
- ▶ Lexica
  - Mauritian Database of inflected verbs compiled on the basis of Carpooran (2009)
  - French **Lexique 3** (New et al., 2007): database of French inflected words with frequency data compiled from post-1950 novels + film subtitles
- ▶ Corpora
  - Written EP **CETEMPúblico** (Santos and Rocha, 2001): tagged corpus of Portuguese (180M words), taken from issues of the newspaper *Público* from 1991 to 1998.
  - Written French 2 years of the newspaper *Le Monde* (2003–2004; 38.5M words), tagged and lemmatized using MElt (Denis and Sagot, 2009)
  - Spoken EP and French **C-ORAL-ROM** (Cresti et al., 2004), collection of balanced corpora of spoken French, Italian, Spanish and Portuguese (~ 300000 words for each language), transcribed, tagged and lemmatized

# Outline

Introduction

Comparing lexifier to creole  
Indo-Portuguese  
Mauritian

Towards explaining the differences

The origin of Mauritian verb forms

Conclusions

References

# The Portuguese conjugation system

- ▶ Portuguese verbal paradigm: 66 cells
- ▶ 3 conjugation classes, each with its own perceptible theme vowel
  - ▶ lavar 'wash' (class1)

TAM	1SG	2SG	3SG	1PL	2PL	3PL
IND.PRS	lav- <b>o</b>	lav <b>a</b> -s	lava <b>a</b>	lav <b>a</b> -mos	lav <b>a</b> -is	lav <b>a</b> -m
IND.PST.IPFV	lav <b>a</b> -va	lav <b>a</b> -vas	lav <b>a</b> -va	lav <b>á</b> -vamos	lav <b>a</b> -veis	lav <b>a</b> -vam
IND.PST.PFV	lav- <b>ei</b>	lav <b>a</b> -ste	lav <b>ou</b>	lav <b>á</b> -mos	lav <b>a</b> -stes	lav <b>a</b> -ram
IND.PST.PRF	lav <b>a</b> -ra	lav <b>a</b> -ras	lav <b>a</b> -ra	lav <b>á</b> -ramos	lav <b>á</b> -reis	lav <b>a</b> -ram
IND.FUT	lav <b>a</b> -rei	lav <b>a</b> -rás	lav <b>a</b> -rá	lav <b>a</b> -remos	lav <b>a</b> -reis	lav <b>a</b> -rão
SBJV.PRS	lav- <b>e</b>	lav <b>e</b> -s	lav <b>e</b>	lav <b>e</b> -mos	lav <b>e</b> -is	lav <b>e</b> -m
SBJV.PST	lav <b>a</b> -sse	lav <b>a</b> -ssem	lav <b>a</b> -sse	lav <b>á</b> -ssemos	lav <b>a</b> -ssem	lav <b>a</b> -ssem
SBJV.FUT	lav <b>a</b> -r	lav <b>a</b> -res	lav <b>a</b> -r	lav <b>a</b> -rmos	lav <b>a</b> -rdes	lav <b>a</b> -rem
COND	lav <b>a</b> -ria	lav <b>a</b> -rias	lav <b>a</b> -ria	lav <b>a</b> -ríamos	lav <b>a</b> -ríeis	lav <b>a</b> -riam
IMP	---	lav <b>a</b>	lav <b>e</b>	lav <b>e</b> -mos	lav <b>a</b> -i	lav <b>e</b> -m
INF.PERS	lav <b>a</b> -r	lav <b>a</b> -res	lav <b>a</b> -r	lav <b>a</b> -rmos	lav <b>a</b> -rdes	lav <b>a</b> -rem
	INF.IMPERS lav <b>a</b> -r		PTCP lav <b>a</b> -do/a		GER lav <b>a</b> -ndo	

# The Portuguese conjugation system

## ▶ lavar 'wash' (class1)

TAM	1SG	2SG	3SG	1PL	2PL	3PL
IND.PRS	lav- <b>o</b>	lav <b>a</b> -s	lav <b>a</b>	lav <b>a</b> -mos	lav <b>a</b> -is	lav <b>a</b> -m
IND.FUT	lav <b>a</b> -rá	lav <b>a</b> -rás	lav <b>a</b> -rá	lav <b>a</b> -remos	lav <b>a</b> -reis	lav <b>a</b> -rão

## ▶ beber 'drink' (class2)

TAM	1SG	2SG	3SG	1PL	2PL	3PL
IND.PRS	beb- <b>o</b>	bebe <b>e</b> -s	bebe <b>e</b>	bebe <b>e</b> -mos	bebe <b>e</b> -is	bebe <b>e</b> -m
IND.FUT	bebe <b>e</b> -rá	bebe <b>e</b> -rás	bebe <b>e</b> -rá	bebe <b>e</b> -remos	bebe <b>e</b> -reis	bebe <b>e</b> -rão

## ▶ subir 'go up' (class3)

TAM	1SG	2SG	3SG	1PL	2PL	3PL
IND.PRS	sub- <b>o</b>	sobe <b>e</b> -s	sobe <b>e</b>	subi <b>i</b> -mos	subi <b>i</b> -s	sobe <b>e</b> -m
IND.FUT	subi <b>i</b> -rá	subi <b>i</b> -rás	subi <b>i</b> -rá	subi <b>i</b> -remos	subi <b>i</b> -reis	subi <b>i</b> -rão

# The Daman / Korlai conjugation system

- ▶ Verbal paradigms in Daman & Korlai creoles: 4 cells
  - ☞ Inflection classes marked by theme vowels
  - ☞ Extension of a 4th class for loans of substratic origin.

	kanta 'sing'	kume 'eat'	subi 'go up'	beblu 'mutter'
BASE	kant <b>a</b>	kum <b>e</b>	sub <b>i</b>	bebl <b>u</b>
PAST	kant-o	kum <b>e</b> -u	sub <b>i</b> -u	bebl <b>u</b>
PROGRESSIVE	kant <b>a</b> -n	kum <b>e</b> -n	sub <b>i</b> -n	bebl <b>u</b> -n
COMPLETIVE	kant <b>a</b> -d	kum <b>i</b> -d	sub <b>i</b> -d	bebl <b>u</b> -d

Daman Creole Portuguese

(adapted from (Clements and Koontz-Garboden, 2002))



# The origin of D/K paradigms

- Each paradigm cell has a clearly identifiable precedent in Portuguese, both in terms of form and in terms of function.

Daman	Portuguese	Daman	Portuguese								
<table border="1"> <tr><td>BASE FORM</td></tr> <tr><td><b>lava</b></td></tr> <tr><td><b>kume</b></td></tr> <tr><td><b>subi</b></td></tr> </table>	BASE FORM	<b>lava</b>	<b>kume</b>	<b>subi</b>	←	<table border="1"> <tr><td>INFINITIVE</td></tr> <tr><td><b>lava-r</b></td></tr> <tr><td><b>come-r</b></td></tr> <tr><td><b>subi-r</b></td></tr> </table>	INFINITIVE	<b>lava-r</b>	<b>come-r</b>	<b>subi-r</b>	
BASE FORM											
<b>lava</b>											
<b>kume</b>											
<b>subi</b>											
INFINITIVE											
<b>lava-r</b>											
<b>come-r</b>											
<b>subi-r</b>											
<table border="1"> <tr><td>COMPLETIVE</td></tr> <tr><td><b>lava-d</b></td></tr> <tr><td><b>kumi-d</b></td></tr> <tr><td><b>subi-d</b></td></tr> </table>	COMPLETIVE	<b>lava-d</b>	<b>kumi-d</b>	<b>subi-d</b>	←	<table border="1"> <tr><td>PST.PTCP</td></tr> <tr><td><b>lava-do/a</b></td></tr> <tr><td><b>comi-do/a</b></td></tr> <tr><td><b>subi-do/a</b></td></tr> </table>	PST.PTCP	<b>lava-do/a</b>	<b>comi-do/a</b>	<b>subi-do/a</b>	
COMPLETIVE											
<b>lava-d</b>											
<b>kumi-d</b>											
<b>subi-d</b>											
PST.PTCP											
<b>lava-do/a</b>											
<b>comi-do/a</b>											
<b>subi-do/a</b>											
<table border="1"> <tr><td>PAST FORM</td></tr> <tr><td><b>lav-o</b></td></tr> <tr><td><b>kume-u</b></td></tr> <tr><td><b>subi-u</b></td></tr> </table>	PAST FORM	<b>lav-o</b>	<b>kume-u</b>	<b>subi-u</b>	←	<table border="1"> <tr><td>PST.PFV</td></tr> <tr><td><b>lav-ou</b></td></tr> <tr><td><b>come-u</b></td></tr> <tr><td><b>subi-u</b></td></tr> </table>	PST.PFV	<b>lav-ou</b>	<b>come-u</b>	<b>subi-u</b>	
PAST FORM											
<b>lav-o</b>											
<b>kume-u</b>											
<b>subi-u</b>											
PST.PFV											
<b>lav-ou</b>											
<b>come-u</b>											
<b>subi-u</b>											
<table border="1"> <tr><td>PROGRESSIVE</td></tr> <tr><td><b>lava-n</b></td></tr> <tr><td><b>kume-n</b></td></tr> <tr><td><b>subi-n</b></td></tr> </table>	PROGRESSIVE	<b>lava-n</b>	<b>kume-n</b>	<b>subi-n</b>	←	<table border="1"> <tr><td>GERUND</td></tr> <tr><td><b>lava-ndo</b></td></tr> <tr><td><b>come-ndo</b></td></tr> <tr><td><b>subi-ndo</b></td></tr> </table>	GERUND	<b>lava-ndo</b>	<b>come-ndo</b>	<b>subi-ndo</b>	
PROGRESSIVE											
<b>lava-n</b>											
<b>kume-n</b>											
<b>subi-n</b>											
GERUND											
<b>lava-ndo</b>											
<b>come-ndo</b>											
<b>subi-ndo</b>											

# Outline

Introduction

Comparing lexifier to creole

Indo-Portuguese

**Mauritian**

Towards explaining the differences

The origin of Mauritian verb forms

Conclusions

References

# The French conjugation system

- ▶ 51 cells

☞ *laver* 'wash':

## Finite forms

TAM	1SG	2SG	3SG	1PL	2PL	3PL
PRS.IND	lav	lav	lav	lav- <b>ɔ̃</b>	lav- <b>e</b>	lav
PST.IND.IPFV	lav- <b>ɛ</b>	lav- <b>ɛ</b>	lav- <b>ɛ</b>	lav-j- <b>ɔ̃</b>	lav-j- <b>e</b>	lav- <b>ɛ</b>
PST.PFV	lave	lava	lava	lava- <b>m</b>	lava- <b>t</b>	lavɛ- <b>ɛ</b>
FUT.IND	lavə- <b>ɛ</b> - <b>ɛ</b>	lavə- <b>ɛ</b> - <b>a</b>	lavə- <b>ɛ</b> - <b>a</b>	lavə- <b>ɛ</b> - <b>ɔ̃</b>	lavə- <b>ɛ</b> - <b>e</b>	lavə- <b>ɛ</b> - <b>ɔ̃</b>
PRS.SBJV	lav	lav	lav	lav-j- <b>ɔ̃</b>	lav-j- <b>e</b>	lav
PST.SBJV	lava- <b>s</b>	lava- <b>s</b>	lava	lava- <b>s</b> -j- <b>ɔ̃</b>	lava- <b>s</b> -j- <b>e</b>	lava- <b>s</b>
COND	lavə- <b>ɛ</b> - <b>ɛ</b>	lavə- <b>ɛ</b> - <b>ɛ</b>	lavə- <b>ɛ</b> - <b>ɛ</b>	lavə- <b>ɛ</b> -j- <b>ɔ̃</b>	lavə- <b>ɛ</b> -j- <b>e</b>	lavə- <b>ɛ</b> - <b>ɛ</b>
IMP	---	lav	---	lav- <b>ɔ̃</b>	lav- <b>e</b>	---

## Nonfinite forms

INF	PRS.PTCP	PST.PTCP			
		M.SG	F.SG	M.PL	F.PL
lave	lav- <b>ũ</b>	lave	lave	lave	lave

# The French conjugation system

- ▶ One productive conjugation (LAVER)
- ▶ Stable but closed second conjugation (FINIR)
- ▶ 61 patterns with 1 to 50 verbs

Conjugation	1 LAVER	2 FINIR	3a RENDRE	3b TENIR	3c CUIRE	3d PEINDRE	3e METTRE	...
Types (Lexique)	5678	282	50	28	28	27	15	
INF	lave	finiʁ	ʁãdʁ	təniʁ	kɥiʁ	pẽdʁ	metʁ	
PST.PTCP	lave	fini	ʁãdy	təny	kɥi	pẽ	mi	
PRS.1SG	lav	fini	ʁã	tjẽ	kɥi	pẽ	me	
PRS.2SG	lav	fini	ʁã	tjẽ	kɥi	pẽ	me	
PRS.3SG	lav	fini	ʁã	tjẽ	kɥi	pẽ	me	
PRS.1PL	lavõ	finisõ	ʁãdõ	təndõ	kɥizõ	pepõ	metõ	
PRS.2PL	lave	finise	ʁãde	təne	kɥize	pepe	mete	
PRS.3PL	lav	finis	ʁãd	tjẽn	kɥiz	pep	met	

# The Mauritian conjugation system

- ▶ The Mauritian verbal paradigm : 2 cells
  - ☞ It distinguishes **morphologically** between long and short forms (Veenstra, 2004; Henri, 2010)
  - ☞ Morphological alternation, (contra Corne, 1982): the alternation is not phonologically predictable

LF	bʷize	bʷije	vãde	amãde	kõsiste	ʷeste	fini	vini
SF	bʷiz	bʷije	van	amãd	kõsiste	ʷes	fini	vin
TRANS.	'break'	'mix'	'sell'	'amend'	'consist'	'stay'	'finish'	'come'

- ▶ The alternation codes syntactic, morphological and/or information-structure oppositions (Henri, 2010)
  - ☞ Here: presence of a nonclausal following complement

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

# Origin of Mauritian paradigms

- ▶ The origin of the forms filling the paradigms is uncertain
  - ▶ Prevalence of syncretism in the French system

Paradigm cells	class 1 LAVER	class 2 FINIR
PRS/IMP.2PL	-e	-ise
IPFV.SG/3PL		
INF		-i
PST.PTCP	∅	
PRS.SG		
PRS.3PL		-is
SBJV.SG/3PL		

- ☞ In 18th century French, infinitive final **-r** was consistently dropped for verbs of all conjugations, except those with a final schwa (Rosset, 1911, Y.-C. Morin, p.c.).
- ▶ Lack of statistically usable historical data (Baker et al., 2007)
- ▶ No parallelism of function

## Interim conclusion

- Conclusion: Indo-Portuguese is more similar to Portuguese than Mauritian is to French

	K/D	Mauritian
survival of inflection class system	yes	no
survival of function of paradigm cells	yes	no
origin of forms	clear	unclear

# Outline

Introduction

Comparing lexifier to creole  
Indo-Portuguese  
Mauritian

Towards explaining the differences

The origin of Mauritian verb forms

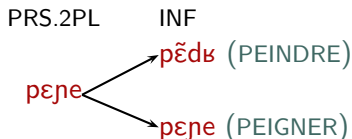
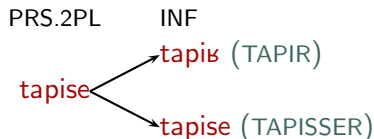
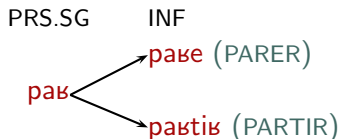
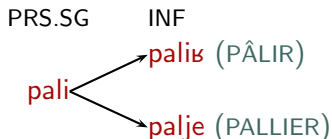
Conclusions

References



# Paradigmatic opacity

- ▶ A form is **paradigmatically opaque** when it is compatible with more than one inflection class.
- ▶ Opaque forms are commonplace in French:



- ▶ Portuguese does not seem to have any paradigmatically opaque forms.

## Comparing paradigm opacity

- ▶ In EP, almost all cells in the paradigm contain a theme vowel precluding paradigm opacity
  - ☞ This is true for all but **1** (PRS.IND.1SG) of the **66** paradigm cells
- ▶ By contrast, in French, only the infinitive, the past participle, the simple past and the (barely used) past subjunctive contain a theme vowel giving unambiguous information on conjugation class.
  - ☞ That is, only **14** out of **51** cells are diagnostic.

	Type frequency	Token frequency (written corpora)	Token frequency (spoken corpora)
Portuguese	98%	99.96%	92.57%
French	27%	33.77%	28.53%

Proportion of paradigm cells with a diagnostic vowel alternation  
(data from **CETEMPúblico**, **Le Monde** and **C-ORAL-ROM**)

## Class-specific syncretism

- ▶ French conjugation is characterized by the high prevalence of inflection class specific patterns of syncretism

	LAVER	FINIR	RENDRE	CUIRE	POUVOIR	DIRE	...
PRS/IMP.2PL	-e	-ise	rãde	kɥize	puve	dit	
IPFV.SG/3PL						dize	
INF	-i		rãdɤ	kɥiɤ	puvwɤ	diɤ	
PST.PTCP							
PRS.SG	∅		vãdy	kɥi	py	di	
PRS.3PL						pø	
SBJV.SG/3PL		-is	-vãd	-kɥiz	pœv	diz	
					pɥis		

- ▶ Portuguese doesn't have **any** class-specific syncretism
- ▶ **Conclusions**
  - ▶ Both paradigm opacity and class-specific syncretism contribute to making French conjugation highly unpredictable
  - ▶ Hence, it would be surprising for a French-based creole to maintain the conjugation system of its lexifier.

## Reorganizing French paradigms

- ▶ On the other hand, the syncretic patterns of French first conjugation verbs are very perceptible

	among 1st conj. tokens		among all verb tokens	
	C-ORAL-ROM	lexique 3	C-ORAL-ROM	lexique 3
'long form'	49.4%	49.1%	14.6%	19.3%
'short form'	40%	40.1%	11.8%	15.8%
contrasting forms	89.4%	89.2%	26.4%	35.2%

### Visibility of the long/short alternation in French

- ▶ If creole formation is at all sensitive to statistical properties of the lexifier's lexicon, this distinction is expected to be present in French-based creoles
- ▶ However since the two forms are highly syncretic, there is no stable function for the form alternants to inherit.

☞ Cf. also Becker and Veenstra (2003); Veenstra (2004)

# Outline

Introduction

Comparing lexifier to creole  
Indo-Portuguese  
Mauritian

Towards explaining the differences

The origin of Mauritian verb forms

Conclusions

References

## Paradigm shape vs. forms in the paradigm

- ▶ Given this story, remaining question: how did the Mauritian **forms** stem from French forms?
  - ▶ Most natural hypothesis: two forms were inherited from French, most likely the **infinitive** and the **present singular**
  - ▶ Alternative hypothesis: one form was inherited from French (most likely the **infinitive**), the alternation itself is native to Mauritian
  - ▶ For the vast majority of verbs the two hypotheses are undistinguishable
- ☞ We checked exhaustively the 1932 verbs whose etymon is undisputably a French verb in Carpooran (2009)

# Possible sources of long and short forms

- Possible sources of the long form:

French cells	example	#
INF or PST.PTCP or . . .	laver > lave	1767
INF	croire > krwar	129
PST.PTCP	offert > ofer	11
PRS.SG	doit > dwa	9
other cases	asseoir > asize	16

- Where the short form coincides with the French PRS.SG:

alternation	examples	#
LF = SF + e	lave	1353
LF = SF	fini, dwa	121
other	—	0

☞ No evidence of direct inheritance of short forms from French

## Remarkable cases

- ▶ In all cases where French has an alternation other than  $LF = SF + e$ , Mauritian uses only one of the two forms.

INF	French	Mauritian		trans.
	PRS.SG	LF	SF	
ale	<b>va</b>	ale	<b>al</b>	'go'
sɔʁtiʁ	<b>sɔʁ</b>	soʁti	soʁt	'exit'
vəniʁ	<b>vjɛ̃</b>	vini	<b>vin</b>	'come'
dəvwɑʁ	dwa	<b>dwa</b>	dwa	'must'
valwɑʁ	vo	<b>vo</b>	vo	'be worth'

- ▶ This is despite the existence of 129 Mauritian verbs whose French etymon **does** have a relevant alternation.
- ▶ Conclusion:
  - ▶ Mauritian verbs stem from a single inherited French form
  - ▶ Even if its shape is influenced by statistical properties of the French lexicon, the alternation is native to Mauritian, not an effect of inheritance of forms



# Outline

Introduction

Comparing lexifier to creole  
Indo-Portuguese  
Mauritian

Towards explaining the differences

The origin of Mauritian verb forms

**Conclusions**

References

# Conclusions

- ▶ We have shown that:
  - ▶ The kind of morphology found in the creoles correlates with highly perceptible properties of the inflectional morphology of the lexifier
    - ▶ For Daman/Korlai : prevalence of theme vowels in Portuguese
    - ▶ For Mauritian : prevalence of syncretism in French
  - ☞ Assumption that the statistical distribution of inflected forms in spoken French and Portuguese have been rather stable over time.
  - ▶ Origin of Mauritian **forms**:
    - ▶ The Mauritian paradigm **always** stems from a **single** French form
    - ▶ That form is **most often** but **not always** the infinitive.
- ▶ Puzzling result: French alternations play a crucial role in shaping Mauritian paradigms, but the alternating **forms** did not survive.
- ▶ Next step: full comparison of paradigm predictability in lexifiers and creoles on the basis of information-theoretic methods (Ackerman et al., 2009; Bonami et al., 2011)
  - ☞ Requires new resources on Portuguese and Indo-Portuguese creoles

- Ackerman, F., Blevins, J. P., and Malouf, R. (2009). 'Parts and wholes: implicative patterns in inflectional paradigms'. In J. P. Blevins and J. Blevins (eds.), *Analogy in Grammar*. Oxford: Oxford University Press, 54–82.
- Baker, P., Fon-Sing, G., and Hookoomsing, V. Y. (2007). 'The corpus of Mauritian Creole texts'. In P. Baker and G. Fon-Sing (eds.), *The making of Mauritian Creole*. London: Battlebridge Publications, 1–61.
- Becker, A. and Veenstra, T. (2003). 'The survival of inflectional morphology in French-related Creoles.' *SSLA*, 25:285–306.
- Bonami, O., Boyé, G., and Henri, F. (2011). 'Measuring inflectional complexity: French and mauritian'. In *Quantitative Measures in Morphology and Morphological Development*. San Diego: University of California.
- Carpooran, A. (2009). *Diksioner Morisien*. Sainte Croix (Mauritius): Koleksion Text Kreol.
- Clements, J. C. (1996). *The Genesis of a Language: The Formation and Development of Korlai Portuguese*. John Benjamins Publishing Company.
- Clements, J. C. and Koontz-Garboden, A. J. (2002). 'Two indo-portuguese creoles in contrast'. *Journal of Pidgin and Creole Languages*, 2:191–236.
- Corne, C. (1982). 'The predicate in Isle de France Creole.' In P. Baker and C. Corne (eds.), *Isle de France Creole. Affinities and Origins*. Ann Arbor: Karoma, 31–48.
- Cresti, E., do Nascimento, F. B., Sandoval, A. M., Veronis, J., Martin, P., and Choukri, K. (2004). 'The C-ORAL-ROM CORPUS. a multilingual resource of spontaneous speech for Romance languages.' In *Proceedings of LREC 2004*.
- Denis, P. and Sagot, B. (2009). 'Coupling an annotated corpus and a morphosyntactic lexicon for state-of-the-art POS tagging with less human effort'. In *Proceedings of PACLIC 2009*.
- Henri, F. (2010). *A Constraint-Based Approach to verbal constructions in Mauritian*. Ph.D. thesis, University of Mauritius and Université Paris Diderot.
- New, B., Brysbaert, M., Veronis, J., and Pallier, C. (2007). 'The use of film subtitles to estimate word frequencies'. *Applied Psycholinguistics*, 28:661–677.
- Rosset, T. (1911). *Les origines de la prononciation moderne étudiées au XVII<sup>e</sup> siècle*. Paris: Armand Colin.
- Santos, D. and Rocha, P. (2001). 'Evaluating cetempúblico, a free resource for Portuguese'. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*. 442–449.
- Veenstra, T. (2004). 'What verbal morphology can tell us about Creole genesis: the case of French-related Creoles'. In I. Plag (ed.), *Phonology and Morphology of Creole Languages*, no. 478 in *Linguistische Arbeiten*. Max Niemeyer Verlag GmbH.

## Sources of long forms: verbs in -i and -e

Alternation	Source of long form	examples	#
$X_e \sim X$	INF/PST.PTCP/...	laver > lave	1353
	non-INF	kone	2
	form+e	assis > asize	2
	ambiguous	mettez/mett- > mete	8
$X_e \sim Y$	INF/PST.PTCP/...	rester > reste	19
$X_e \sim X_e$	INF/PST.PTCP/...	jongler > zongle	284
	other	né > ne	3
$X_i \sim X_i$	INF/PST.PTCP/...	fini > fini	109
$X_i \sim X$	INF/PST.PTCP/...	sortir > sorti	2

Mauritian verbs ending in -e or -i

## Sources of long forms: other verbs

- ▶ The remaining verbs are all non-alternating

French paradigm cell	examples	#
INFINITIVE	dire > dir vivre > viv croire > krwar	129
PST.PTCP	mort > mor foutu > fouti offert > ofer	11
PRS.SG	doit > dwa connait > kone vaut > vo	9
Other	déteint > detin éclos/éclot > eklo ?éteigne > tengn	3

Mauritian verbs not ending in -e or -i