

Investigating a discriminative approach to creolization

Fabiola Henri

University of Kentucky
fshe223@uky.edu

Creoles: A sociohistorical definition

- ▶ Creole languages are new vernaculars that resulted from contact between colonial non-standard varieties of a European language and several non-European languages on colonial European plantation settlements around the Atlantic, the Indian and Pacific Oceans between the 16th and 19th century.
- ▶ 93 creoles on Ethnologue: varieties based on French, English, Portuguese, Spanish, Dutch, German.
 - ▶ Extension to (1) varieties that developed from contact between non-European languages (2) pidgins that are native languages

The structural perspective

Our prototype is a creole with no inflectional affixes, no use of tone to contrast monosyllables or encode syntax, and derivational affixes whose semantic contribution is consistently transparent [...] these traits are not an arbitrary conglomerate, but the direct result of severely interrupted transmission of a lexifier, at too recent a date for the traits to have been undone by diachronic change.

MCWHORTER 1998

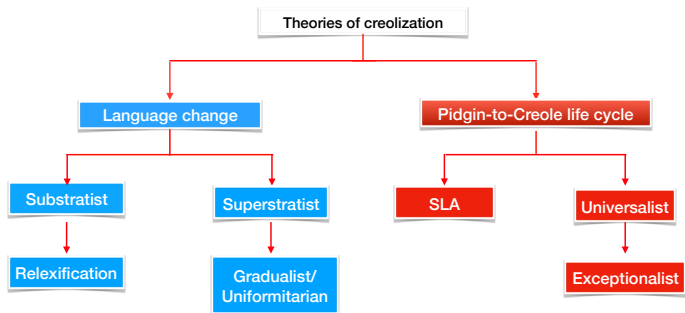
Creole languages result from the adaptation of a language, especially some Indo-European language, to the (so to speak) phonetic and grammatical genius of a race that is linguistically inferior.

VINSON 1889

Le maniement du verbe français avec ses flexions de mode, de temps, de nombre et de personne, offrait des complications que le créole devait nécessairement écarter. Ici la simplification a été poussée à ses dernières limites. Le thème verbal n'a qu'une forme unique : mo vini je viens; to té vini tu es venu; li va vini il viendra; etc., etc.

BAISSAC 1880

Theoretical bases



Morphology as diagnostic

- ▶ Post-Bloomfieldian American structuralism perspective on word structure.

French at most 3 suffixes

Mauritian no true affixation

- | | |
|----------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| <p>(1) a. all-ons
go[PRS]-1PL</p> <p>b. i-r-i-ons
go-FUT-ANA-1PL</p> <p>c. va-s
go[PRS]-2SG</p> | <p>(2) a. ale
go-LF</p> <p>b. al
go-SF</p> <p>c. va
IND.IRR</p> |
|----------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|

- ▶ Claims are based on hand-picked samples except for recent work by Roberts and Bresnan (2008); Bakker et al. (2011) and notably Blasi et al. (2017)

Word and Pattern Morphology

- ▶ Fusional languages are not amenable to the agglutinative ideal (Matthews, 1972)
- ▶ Two dimensions of the **word** in morphological analysis (e.g. Ackerman et al., 2009; Blevins et al., 2016; Blevins, 2016)
 - Words are stable units that sanction reliable deductions about other words.
 - Surface patterns of words cohere into organized systems.
- ▶ Implicative WP: Complex system of interdependencies crucial to language learning and language change.
 - Patterns of featuring variable contrasts persist throughout language change (Maiden, 2004, 2018)

Morphomic patterns

- Morphomic patterns are phonologically and/or functionally unmotivated patterns of alternation (Maiden, 2004, 2018)
- These idiosyncratic patterns may arise through sound change or from old morphomic patterns

	1SG	2SG	3SG	1PL	2PL	3PL
PRS.IND	bwa	bwa	bwa	buv-õ	buv-e	bwav
IMP		bwa		buv-õ	buv-e	
PST.IND.IPFV	buv-ε	buv-ε	buv-ε	buv-jõ	buv-je	buv-ε

Partial paradigm of French verb boire ‘drink’

Outline

Introduction

Recurrent Patterns

Verbs

Marking

Nouns

The discriminative approach

Concluding remarks

Root-based

► 51 cells

☞ laver ‘wash’:

Finite forms						
TAM	1SG	2SG	3SG	1PL	2PL	3PL
PRS.IND	lav-e	lav-es	lav-e	lav-ons	lav-ez	lav-ent
PST.IND.IPFV	lav-ai-s	lav-ai-s	lav-ai-t	lav-i-ons	lav-i-ez	lav-ai-ent
PST.PFV	lav-ai	lav-as	lav-a	lav-â-mes	lav-â-tes	lav-èr-ent
FUT.IND	lav-er-ai	lav-er-as	lav-er-a	lav-er-ons	lav-er-ez	lav-er-ons
PRS.SBJV	lav-e	lav-es	lav-e	lav-i-ons	lav-i-ez	lav-ent
PST.SBJV	lav-ass-e	lav-ass-es	lav-ât	lav-ass-i-ons	lav-ass-i-ez	lav-ass-ent
COND	lav-er-ais	lav-er-ais	lav-er-ait	lav-er-i-ons	lav-er-i-ez	lav-er-aient
IMP	---	lav-e	---	lav-ons	lav-ez	---

Nonfinite forms					
INF	PRS.PTCP	PST.PTCP			
		M.SG	F.SG	M.PL	F.PL
lav-er	lav-ant	lav-é	lav-ée	lav-és	lav-ées

Stem-based

► 51 cells

☞ laver ‘wash’:

Finite forms

TAM	1SG	2SG	3SG	1PL	2PL	3PL
PRS.IND	lav	lav	lav	lav- õ	lav- e	lav
PST.IND.IPFV	lav- ε	lav- ε	lav- ε	lav-j- õ	lav-j- e	lav- ε
PST.PFV	lavε	lava	lava	lava- m	lava- t	lavε- ɸ
FUT.IND	lavə- ɸ-ε	lavə- ɸ-a	lavə- ɸ-a	lavə- ɸ-õ	lavə- ɸ-e	lavə- ɸ-õ
PRS.SBJV	lav	lav	lav	lav-j- õ	lav-j- e	lav
PST.SBJV	lava- s	lava- s	lava	lava- s-j-õ	lava- s-j-e	lava- s
COND	lavə- ɸ-ε	lavə- ɸ-ε	lavə- ɸ-ε	lavə- ɸ-j-õ	lavə- ɸ-j-e	lavə- ɸ-ε
IMP	---	lav	---	lav- õ	lav- e	---

Nonfinite forms

INF	PRS.PTCP	PST.PTCP			
		M.SG	F.SG	M.PL	F.PL
lave	lav- ã	lave	lave	lave	lave

Word-based

► 51 cells

👉 laver ‘wash’

Finite forms						
TAM	1SG	2SG	3SG	1PL	2PL	3PL
PRS.IND	lav	lav	lav	lavǫ	lave	lav
PST.IND.IPFV	lave	lave	lave	lavjǫ	lavje	lave
PST.PFV	lave	lava	lava	lavam	lavat	lavεε
FUT.IND	ve lave	va lave	va lave	alǫ lave	ale lave	vǫ lave
PRS.SBJV	lav	lav	lav	lavjǫ	lavje	lav
PST.SBJV	lavas	lavas	lava	lavasjǫ	lavasje	lavas
COND	lavøεe	lavøεe	lavøεe	lavøεjǫ	lavøεje	lavøεe
IMP	---	lav	---	lavǫ	lave	---

Nonfinite forms					
INF	PRS.PTCP	PST.PTCP			
		M.SG	F.SG	M.PL	F.PL
lave	lavǫ	lave	lave	lave	lave

The French conjugation system

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

	LAVER	AMENER	FINIR	RENDRE	CUIRE	POUVOIR	DIRE	...
Type Frequency	4193	89	581	10	24	1	1	...
PRS/IMP.2PL	lave	aməne	finise	rāde	kʷize	puve	dit	...
IPFV.SG/3PL			fini	rādɛ	kʷiɛ	puvvaɛ	dize	...
INF				vādɥ	kʷi	py	diɛ	...
PST.PTCP	lav	amɛn	finis	vā	kʷiz	pø	di	...
PRS.SG				-vād	-kʷiz	pœv	diz	...
PRS.3PL						pʷis		...
SBJV.SG/3PL								...

- ☞ In 18th century French, infinitive final **-r** was consistently dropped for verbs of all conjugations, except those with a final schwa (Y.-C. Morin, p.c.)

From French to Creole: Mauritian

The Mauritian verbal system: 2 cells paradigm

- Lexical database of 2039 verbs (based on Carpooran (2011)'s monolingual dictionary)
- It distinguishes **morphologically** between long and short forms (Veenstra, 2004; Henri, 2010)
- 70% with distinct LF and SFs ; 30% syncretic

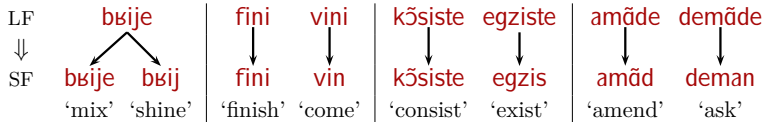
LF	kɔije	bɔije	luke	vāde	amāde	paste	ɤeste	teste	fini	vini	soʔti	ofeʔ	tʃombo
SF	kɔij	bɔije	luk	van	amād	pas	ɤes	teste	fini	vin	soʔt	ofeʔ	tʃom

TRANS. 'shout' 'mix' 'peep' 'sell' 'amend' 'filter' 'stay' 'test' 'finish' 'come' 'go out' 'offer' 'hold'

The Mauritian conjugation system

Against phonologizing accounts

- ▶ Syncretic forms are as productive and regular as e-final forms
- ▶ Neither form is uniformly predictable from the other (pace Corne, 1982):



Reorganizing French paradigms

The role of analogy

3 conjugation groups: **e** final, **i** final , **X** final

- Productive groups: **e** and **X** final
- (New) verbs can either integrate the system in the 1st or 3rd conjugation
- The cell in which they initially appear is unpredictable
- The crucial role of analogy and levelling

Language	Mauritian		trans.
	LF	SF	
Fr. PST.PART.FEM asiz ‘sit’	asize	asiz	‘sit’
Fr. PRS.SG kone ‘know’	kone	kon	‘know’
Fr. kōdijis ‘drive’	kōdire	kōdir	‘drive’
Eng. luk ‘look’	luke	luk	‘peep/look’
Eng. buk ‘book’	buk	buk	‘book’
Creole	mulugāde	mulugan	‘plot’

- ☞ Class assignment is not based on form (eg. **kōdire** vs **prodire**, **luke** vs **buk**)
- ☞ Class assignment is not based on origin of conjugation class in Fr. (eg. **ale** ~ **al**, **vande** ~ **vann** ‘sell’)

Morphologization of the alternation

This development parallels the evolution of the inflectional system from Latin to Romance

- Phonological deformation:
consonant cluster
reduction
- Lexical insertions

	French		'trans'
	ɤãde	ɤãn	'give back'
	demãde	deman	'ask'
	ɤeste	ɤes	'stay'
	tõbe	tõm	'come'
	ɤõfle	ɤõf	'snore'
	ɛgzaʒɛɤe	ɛgzaʒɛɤ	'exaggerate'
	paste	pas	'filter'
	tʃombo	tʃom	'hold'

The French conjugation system

The syncretic patterns of French first conjugation verbs are very perceptible

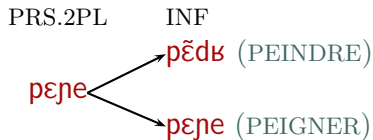
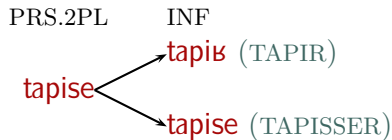
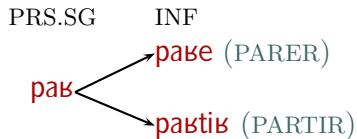
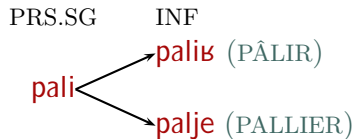
	among 1st conj. tokens		among all verb tokens	
	C-ORAL-ROM	lexique 3	C-ORAL-ROM	lexique 3
'long form'	49.4%	49.1%	14.6%	19.3%
'short form'	40%	40.1%	11.8%	15.8%
contrasting forms	89.4%	89.2%	26.4%	35.2%

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 (cf. also Chaudenson, 2003; Becker and Veenstra, 2003; Veenstra, 2004)
- ☞ Differences in frequency of specific forms also affect perceptibility
- ☞ However since the two forms are highly syncretic, there is no stable function for the form alternants to inherit

Paradigmatic opacity

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



- A paradigm **cell** is opaque if the exponents in contains do not allow to unambiguously predict the lexeme's inflection class.

Exaptation of the function

- The contexts in which the two forms appear do not form natural classes (Henri and Abeillé, 2008; Henri, 2010)

		Distribution	SF	LF
Morphology				
Reduplication		reduplicant	yes	no
		base	yes	yes
Conversion			yes	yes
Syntax				
		V with canonical phrasal complements	yes	no
		V with no complements	no	yes
		V with clausal complements	no	yes
		V with true adjuncts	no	yes
		Extracted complements	no	yes
Information Structure				
Verum Focus	→	V with canonical phrasal complements	no	yes
Object Focus	→	V with applicativized adjuncts	yes	no
Verb Focus	→	V with applicativizable adjuncts	no	yes

Morphomic distribution

Syntax & Information structure

- ▶ The SF appears with nonextracted nonclausal complements
 - No adjacency requirement.

(3) a. Mo **manz** toultan poul.
 1SG eat.SF always chicken
 ‘I always eat chicken.’

- ▶ When the LF appears, it encodes Verum Focus (Henri et al., 2008; Henri, 2010)
 - It is used in counter-oriented moves (deferments, counter-implicative and counter-propositional moves).

(4) a. Mo pa **MANZE** poul!
 1SG PST think Mary NEG eat.LF/ curry chicken
 ‘I DON’T eat chicken!’

Function: Morphomic distribution

Semantically defined adjuncts

- ▶ The SF appear with ‘applicativized’ adjuncts (locative, instrumental, degree, frequency, manner, temporal adjuncts)
 - They are focused and integrated as objects
 - This structure is more canonical in its use (Hassamal & al. 2017).

- (5) a. **Mo manz** ar lame.
 1SG eat.SF PREP 1SG.POSS hand
 ‘I eat WITH MY HANDS.’
- b. **Mo manz** bokou/bien/vit.
 1SG eat.SF/*LF a_lot/well/fast
 ‘I eat a lot/well/fast.’

- ▶ ‘Applicativizable’ adjuncts trigger LFs: The verb is in focus.

- (6) a. **Mo manze** ar lame.
 1SG IRR.DEF eat.SF PREP 1SG.POSS hand
 ‘I EAT with my hands.’
- b. **Mo manze** bokou/bien/vit.
 1SG eat.SF/*LF a_lot/well/fast
 ‘I EAT a lot/well/fast.’

In lexeme-formation: Reduplication

- ▶ The two forms are used in “attenuative” reduplication which is a derivational process creating new verbal lexemes (Henri, 2010)
- ▶ Predictiveness: Speakers predict SFs of strictly intransitive verbs based on reduplication

LF	SF	gloss	red. LF	red. SF	trans.
sāte	sāt	‘sing’	sātsāte	sātsāt	‘hum’
reste	res	‘stay’	ɸesɸeste	ɸesɸes	‘stay occasionally’
soəti	soət	‘get out’	soətsoəti	soətsoət	‘get out occasionally’
balje	balje	‘sweep’	baljebalje	baljebalje	‘sweep carelessly’

Examples of attenuative reduplication

In lexeme-formation: Conversion

- $V \rightarrow N$ conversion may select either a verb's LF or its SF

- (7) a. (i) **dāse**_{LF} 'to dance' > **dāse**_N 'dancing'
 (ii) **dāns**_{SF} 'to dance' > (la)**dās**_N '(definite) dance'
- b. **luke**_{LF} 'to peep' > **luke**_N 'peeping'
- c. **res**_{SF} 'to remain' > (le)**res**_N '(definite) leftover'
- (8) a. **karja**_N 'termite' > **karjat**_{SF}, **karjate**_{LF} 'to infect'
- b. **lasas**_N 'hunting' > **lasas**_{LF/SF} 'to hunt'

Function: Morphomic distribution

Substrate influence

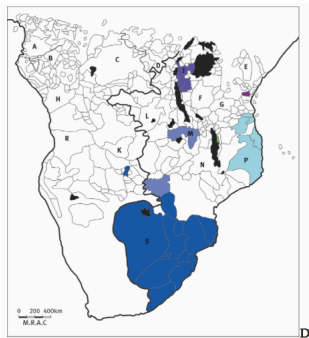


Figure 1: The geographical distribution of the cj/dj alternation

Bantu languages feature a contrast between a SF and a LF — disjoint/conjoint, applicative, tone contrasts, ... (van der Wal and Hyman, 2016)

- The distribution of these forms is not straightforward in Bantu
- Implicational relations that predicts distribution

Function: Morphomic distribution

Substrate influence

- ▶ In many of those languages, SFs never appear in sentence final position
- ▶ These forms are used for structural and focus relations and in reduplications

(9) a. **dikgomó dí-fúla** kwa nokeng.

10.cow SM10-CJ.graze at river

‘The cows graze/are grazing at the river.’

b. **dikgomó dí-á-fúla**

10.cow SM10-PRS.DJ-graze

‘The cows graze/are grazing.’ [Tswana, S31]

(10) a. **thólá** → **oo-thólá-thólá**

search 1.PERF.DJ-search-RED ‘search repeatedly’

[Makhuwa, P31]

b. **hadíno** → **hadinodiíno** [Malagasy (areal)]

‘forget’ ‘forget a bit’

- ☞ Predictiveness: The LF/SF alternation in Mauritian exapts to assume a (morphomic) distribution analogous to the one seen in Bantu

From French to French-based creoles

- ▶ The pattern is found in varying degrees in other FbC with disparate functions
 - The forms do not originate from the same French cells/verb
 - Gua. seems to also have retained PRS.PART forms and Reu. the COND
 - French-based creoles spoken in the Americas do not share substrates with those of the Indian Ocean or the Pacific.

		Form	Function	Nº of distinct alt.
Reu.	LF	ale, konèt, vène, ...	~ Mauritian	~ 70%
	SF	al, kone, vèn, ...		
LC	LF	ale, konen, vini, ...	PST/IMP.2F	~ 50%
	SF	al, konen, vin, ...	PRS/IMP	
Gua.	LF	ale, save, vini, ...	PST.IPFV	34/1825
	SF	ay, sav, vin, ...	PRS.PERF	
Hai	LF	ale, konen, vini, ...	Phonological(?)	12/2657
	SF	al/ay, konen, vini, ...		

- ▶ Lexeme formation: Allomorphic stems are also visible across morphological families e.g. Hai. **babouké** ‘constrain’ → **babouk** ‘constraint’; Mau. **bwar** ‘drink_V’ ~ **labwason** ‘drink_N’ ~ **labivet** ‘bar’ ~ **biver** ‘drinker’

From lexifier to creole

The Portuguese conjugation system

- ▶ lavar ‘wash’ (class1)

TAM	1SG	2SG	3SG	1PL	2PL	3PL
IND.PRS	lav- o	lava- s	lava a	lava- mos	lava- is	lava- m
IND.FUT	lava- r á	lava- r ás	lava- r á	lava- r emos	lava- r eis	lava- r ão

- ▶ beber ‘drink’ (class2)

TAM	1SG	2SG	3SG	1PL	2PL	3PL
IND.PRS	beb- o	bebe- s	bebe e	bebe- mos	bebe- is	bebe- m
IND.FUT	bebe- r á	bebe- r ás	bebe- r á	bebe- r emos	bebe- r eis	bebe- r ão

- ▶ subir ‘go up’ (class3)

TAM	1SG	2SG	3SG	1PL	2PL	3PL
IND.PRS	sub- o	sobe- s	sobe e	subi- mos	subi- s	sobe- m
IND.FUT	subi- r á	subi- r ás	subi- r á	subi- r emos	subi- r eis	subi- r ão

From lexifier to creole

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 (Gujrati, Marathi, ...).

	kanta	kume	subi	beblu
	'sing'	'eat'	'go up'	'mutter'
BASE	kanta	kume	subi	beblu
PAST	kant-o	kume-u	subi-u	beblu
PROGRESSIVE	kanta-n	kume-n	subi-n	beblu-n
COMPLETIVE	kanta-d	kumi-d	subi-d	beblu-d

Daman Creole Portuguese
Adapted from Clements (2002)

From lexifier to creole

Origin of Daman / Korlai conjugation

- ▶ Indo-Portuguese has retained much of the Portuguese structure
 - 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>lava</td></tr> <tr><td>kume</td></tr> <tr><td>subi</td></tr> </table>	BASE FORM	lava	kume	subi	⇐	<table border="1"> <tr><td>INFINITIVE</td></tr> <tr><td>lava-r</td></tr> <tr><td>come-r</td></tr> <tr><td>subi-r</td></tr> </table>	INFINITIVE	lava-r	come-r	subi-r												
BASE FORM																						
lava																						
kume																						
subi																						
INFINITIVE																						
lava-r																						
come-r																						
subi-r																						
		<table border="1"> <tr><td>COMPLETIVE</td></tr> <tr><td>lava-d</td></tr> <tr><td>kumi-d</td></tr> <tr><td>subi-d</td></tr> </table>	COMPLETIVE	lava-d	kumi-d	subi-d	⇐	<table border="1"> <tr><td>PST.PTCP</td></tr> <tr><td>lava-do/a</td></tr> <tr><td>comi-do/a</td></tr> <tr><td>subi-do/a</td></tr> </table>	PST.PTCP	lava-do/a	comi-do/a	subi-do/a										
COMPLETIVE																						
lava-d																						
kumi-d																						
subi-d																						
PST.PTCP																						
lava-do/a																						
comi-do/a																						
subi-do/a																						
<table border="1"> <tr><td>PAST FORM</td></tr> <tr><td>lav-o</td></tr> <tr><td>kume-u</td></tr> <tr><td>subi-u</td></tr> </table>	PAST FORM	lav-o	kume-u	subi-u	⇐	<table border="1"> <tr><td>PST.PFV</td></tr> <tr><td>lav-ou</td></tr> <tr><td>come-u</td></tr> <tr><td>subi-u</td></tr> </table>	PST.PFV	lav-ou	come-u	subi-u		<table border="1"> <tr><td>PROGRESSIVE</td></tr> <tr><td>lava-n</td></tr> <tr><td>kume-n</td></tr> <tr><td>subi-n</td></tr> </table>	PROGRESSIVE	lava-n	kume-n	subi-n	⇐	<table border="1"> <tr><td>GERUND</td></tr> <tr><td>lava-ndo</td></tr> <tr><td>come-ndo</td></tr> <tr><td>subi-ndo</td></tr> </table>	GERUND	lava-ndo	come-ndo	subi-ndo
PAST FORM																						
lav-o																						
kume-u																						
subi-u																						
PST.PFV																						
lav-ou																						
come-u																						
subi-u																						
PROGRESSIVE																						
lava-n																						
kume-n																						
subi-n																						
GERUND																						
lava-ndo																						
come-ndo																						
subi-ndo																						

From lexifier to creole

Other creoles

☞ Organization of the system and frequency of input matters

▶ Other Portuguese-based creoles

- Port. SG.PST.IMPV **kantaβa** → GB Kriyol, Cape Verdian, ... **kantaba** with only clitic pronouns allowed to intervene between the base form and **-ba**

▶ English-based creoles similarly show varying inherited forms

- Sranan: e.g. **wan** ~ **wani** from the Eng. Base and PROG forms
- Other English-based creoles also feature past tense forms including Eng. irregular stems in their paradigm AAVE **fid** ~ **fɛd**
- ...

Outline

Introduction

Recurrent Patterns

Verbs

Marking

Nouns

The discriminative approach

Concluding remarks

Tense, Mood, Aspect

- ▶ The WP approach to morphological structure fully integrates periphrastic constructions
- ▶ The development of periphrastic marking is unsurprising given the donor languages (See also Bybee 1985 and subseq.)
 - English makes heavy use of periphrases to express TMA (syntactic)
 - Similarly in colloquial French e.g. *composé*, the future in **ale** ‘go’ + V, etc.
 - ▶ Abouda and Skrovec (2017) show that the periphrastic future is more widely used in Modern French than its synthetic counterpart.
 - ▶ This is reflected in French varieties such as Acadian French, Cajun French where the paradigm is levelled with only a few distinct synthetic forms; Mood and aspect are expressed periphrastically

- (11) a. **Je suis après manger.** ‘I am eating.’
b. **Je vas manger.** ‘I will eat.’

Tense, Mood, Aspect

	Tense		Mood			Aspect	
		IRR.DEF	IRR.IND	COND	PRF	PROG	
Mauritian	ti	pou	ava	—	finn	pe	
Reunionese	te	pou	ava	lore	fini	pou la apre	
Haitian	te	—	av	—	fini	ap	
Guadeloupean	te	ke	—	—	fini	ka	
Louisiana C.	te	va e	se	se	bin sa	ape	

Partial TMA marking in French-based creoles

- ▶ Intricate distribution — e.g. LC uses the FUT **va** in affirmatives and **e** with negation.
- ▶ Reunionese uses **i** to mark finite verbs.
- ▶ Note that marking does not always appear preverbally (cf. Portuguese-based creoles).

Tense, Mood, Aspect

- Grammaticalization: The markers do not have a lexical counterpart and they are subject to sandhi processes

(12) Mo **tj=a=pj=aprann** **si mo ti kone.**
 1SG.WF PST=IRR.IND=PRF=study.LF if 1SG.WF PST know.LF
 ‘I would have studied if I knew.’ [Mauritian]

- Inflectional periphrases: non compositional meaning, dependent of a verbal host.

(13) a. An **ka vinn** vwè’w.
 1SG PROG come.SF see.3SG
 ‘I am coming to see you.’ (once)

b. An **ka vini** vwè’w.
 1SG PROG come.LF see.3SG
 ‘I will come to see you.’ (possibly multiple times) [Guadeloupean]

Negation

- ▶ Contrary to common belief, sentential negation show variable ordering in many creoles. It is also subject to sandhi.
- ▶ In French, negation follows finite forms but precede non-finite forms.
 - In LC, the distribution of negation correlates with verb form.

- (14) a. Mo **lav** **pa** mo figi toulejou.
 1SG.WF wash.PRST.IND NEG 1SG.POSS face everyday
 ‘I wash my face everyday.’
- b. Mo **pa lave** mo figi toulejou.
 1SG NEG wash.PST 1SG.POSS face everyday
 ‘I didn’t wash my face everyday.’ [Louisiana Creole]

- In Mau., only neg-raising verbs allow the variation. In the presence of TMA markers, negation always precede the verb

- (15) a. Mo **pans pa** ki li pou vini.
 1SG.WF think.SF NEG that 3SG IRR.DEF come.LF
 ‘I think that he won’t come.’
- b. Mo **papanse** ki li pou vini
 1SG NEG wash.PST 1SG.POSS face everyday
 ‘I didn’t wash my face everyday.’ [Mauritian]

Article agglutination

- ▶ Creole words often agglutinate the phonology of two words from the lexifier
 - In French-based Creoles, inherited nouns come from the reanalysis of the sequence **det** + **noun** in the lexifier.

French article		noun	example	trans.	size
la	⊕	tabl	latab	‘table’	457
le	⊕	tub	letuʁ	‘turn’	49
	⊕	li	lili	‘bed’	11
l	⊕	amuʁ	lamuʁ	‘love’	723
dy	⊕	te	dite	‘tea’	37
dəl	⊕	o	delo	‘water’	3
ma	⊕	tāt	matāt	‘aunt’	3
mō	⊕	pεʁ	mōpeʁ	‘father’	1
ind.- n	⊕	εspes	nespes	‘species’	4
plur.- z	⊕	animo	zanimo	‘animal’	62
TOTAL					1350/4760

Sample of Mauritian agglutinations

Predicting agglutinations

- ▶ Factors jointly predicts agglutination to occur in Mauritian (simple logistic regression 88.5% accuracy). (Henri and Bonami, 2017)
 - ▶ Each is a pretty bad predictor on its own

Predictor	Accuracy	Acc. increase	$D_{x,y}$
Monosyllabicity or Polysyllabicity	0.8252101	0	0.180
Initial segment type	0.8382353	0.0130252	0.464
Gender	0.8252101	0	0.274
Frequency of coll. with DEF.SG	0.8252101	0	0.274
Raw frequency	0.8254202	0.0002101	0.349
Age	0.8252101	0	0.277
all	0.8798319	0.0546218	0.846
Baseline (no predictor)	0.8252101		

- ☞ Substratic influence — does **LA + N** makes a better class marker?

Nominal alternations

- ▶ Interestingly, some also show alternating forms

(16) a. Donn mwa enn **liv** pwason.
 give.SF 1SG.STF IND pound fish
 ‘Give me a pound of fish.’

b. Komie ou dir **laliv**?
 how_much 2SG.FOR say.SF pound
 ‘How much do you say the pound?’

[Mauritian]

- ▶ Among the 1240 nouns with an agglutinated definite article, **275** are alternating
 - The function seems to be morphomic
- ▶ In Martiniquais, the agglutinated **LA + N** form is a semantic definite that contrasts with the bare form and nouns marked for specificity **plaj** ~ **laplaj** ~ **plaj-la** ‘(the) beach’ ; **LE + N** marks the plural on proper nouns **leñki** ‘The Henries’

Outline

Introduction

Recurrent Patterns

Verbs

Marking

Nouns

The discriminative approach

Concluding remarks

Implicational relations

- ▶ The survival and exaptation of patterns of relatedness across creoles support the idea of a different internal organization of the word (synthetic and periphrastic)
 - This organization essentially implicational and can be formalized in information-theoretic terms (Ackerman et al., 2009)
 - Measuring the uncertainty of predicting cell X based on knowledge of cell Y using conditional entropy

(Mauritian) LF \mapsto SF	(French) INF \mapsto PRS	(Mauritian) SF \mapsto LF	(French) PRS \mapsto INF
0.563	0.338	0.925	0.355

Bonami & Henri (2010)

- Conclusion: predicting one cell of the paradigm from another **on the basis of morphological information** is noticeably more complex in Mauritian than in French.
- ☞ A different kind of complexity that highlights speakers strategies wrt learning.

Creolization as discriminative learning

- ▶ The relation between form and meaning is essentially discriminative (e.g. Baayen et al., 2015; Blevins, 2016; Baayen et al., 2018)
 - There is no necessity for predefined primitives such as the morpheme despite the fact that these units can be abstracted iff it is useful for the organization of a language
 - Discriminative learning proceeds by contrasts, naturally incorporating sequences involving frequent collocations
 - ▶ cf. TMA and agglutinations
- ▶ The model integrates the importance of proportional analogies

Creolization as discriminative learning

- ▶ Psychological reality: categories and classes are constantly recalibrated through experience
 - Dynamic discriminative learning naturally accounts for semantic drifts or exaptation
 - Exaptation of the pattern of alternation also correlates with the speaker's prior experience wrt. his native language
 - Different linguistic ecologies will yield different outcomes (variety of colonial languages, substratic languages, setting, ...)

Outline

Introduction

Recurrent Patterns

Verbs

Marking

Nouns

The discriminative approach

Concluding remarks

Ramifications for creolization

- ▶ Creoles **have** morphology!
- ▶ Creoles **are** genetically related to **both** their lexifiers and their substrates (see also Aboh, 2015; Blasi et al., 2017).
 - There is no evidence that they arise from ‘pidgins’
- ▶ Their exceptionality lies in the fact that they emerge in a short period of time, out of contact of many languages, often typologically unrelated.
- ▶ From an integrative perspective, creoles can be equally or more complex than the lexifier
- 👉 Proposal: The emergence of creoles is one manifestation of discriminative learning

Resources

- ▶ The necessity of applying quantitative methods in order to identify patterns of organization that are otherwise undetectable
- ▶ Creoles are low-resource languages:
 - Few digitally available data — available raw dataset require orthographic normalization, documentation, ...
 - The sombre history of slavery makes it difficult to clearly identify substratic contribution
 - Scarcity of diachronic data of colonial varieties spoken during slavery — A few old handwritten texts could be OCRed
 - Identifying code-switching
 - Type and frequency of input matter as much as the sociolinguistic ecology
- ▶ Simulations that could produce an outcome similar to creoles?

Mersi!



References

- Aboh, E. O. (2015). *The Emergence of Hybrid Grammars*. Cambridge Approaches to Language Contact. Cambridge University Press.
- Abouda, L. and Skrovec, M. (2017). 'Du rapport micro-diachronique futur simple/futur périphrastique en français moderne. étude des variables temporelles et aspectuelles'. Corela [En ligne].
- 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.
- Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., and Blevins, J. P. (2018). 'The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning'.
- Baayen, R. H., Shaoula, C., Willitsb, J., and Ramscara, M. (2015). 'Comprehension without segmentation: A proof of concept with naive discriminative learning'.
- Bakker, P., Daval-Markussen, A., Parkvall, M., and Plag, I. (2011). 'Creoles are typologically distinct from noncreoles'. *Journal of Pidgin and Creole Languages*, 26:5–42.
- Becker, A. and Veenstra, T. (2003). 'The survival of inflectional morphology in French-related Creoles.' *SSLA*, 25:285–306.
- Blasi, D. E., Michaelis, S. M., and Haspelmath, M. (2017). 'Grammars are robustly transmitted even during the emergence of creole languages'. *JNature Human Behavior*:723–729.
- Blevins, J. P. (2016). *Word and Paradigm Morphology*. Oxford University Press.
- Blevins, J. P., Ackerman, F., and Malouf, R. (2016). 'Word and Paradigm Morphology'. In J. Audring and F. Masini (eds.), *Oxford Handbook of Morphological Theory*. Oxford.
- Carpooan, A. (2011). *Diksjoner Morisien. Sainte Croix (Mauritius): Koleksion Text Kreol*.
- Chaudenson, R. (2003). *La Créolisation: Théorie, Applications, Implications*. Paris: L'Harmattan.
- 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.
- Henri, F. (2010). *A Constraint-Based Approach to verbal constructions in Mauritian*. Ph.D. thesis, University of Mauritius and Université Paris Diderot.
- Henri, F. and Abeillé, A. (2008). 'Verb form alternations in Mauritian'. In S. Müller (ed.), *Proceedings of the 15th Conference on HPSG*. Stanford: CSLI Publications, 378–398.
- Henri, F. and Bonami, O. (2017). 'Prédire l'agglutination de l'article en mauricien'. *Faits de Langues*, 49.
- 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.
- Maiden, M. (2004). 'Morphological autonomy and diachrony'. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2004*. Dordrecht: Kluwer, 137–175.
- (2018). *The Romance Verb: Morphomic Structure and Diachrony*. Oxford: OUP.
- Matthews, P. H. (1972). *Inflectional morphology: a theoretical study based on aspects of Latin verb conjugation*. No. 6 in *Cambridge studies in linguistics*. Cambridge Univ Press.
- Roberts, S. and Bresnan, J. (2008). 'Retained inflectional morphology in pidgins: A typological study'. *Linguistic Typology*, 12:269–302.
- van der Wal, J. and Hyman, L. M. e. (2016). *The Conjoint/Disjoint Alternation in Bantu*. Berlin, Boston: Mouton de Gruyter.
- 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.