

LANGUAGE CHANGE AND LANGUAGE CONTACT
IN PIDGINS AND CREOLES

CREOLE LANGUAGE LIBRARY (CLL)

A companion series to the
“JOURNAL OF PIDGIN & CREOLE LANGUAGES”

Editors

Jacques Arends (Amsterdam)
John Victor Singler (New York)

Editorial Advisory Board

Mervyn Alleyne (Kingston, Jamaica)
Norbert Boretzky (Bochum)
Lawrence Carrington (Trinidad)
Glenn Gilbert (Carbondale, Illinois)
George Huttar (Dallas)
John Holm (New York)

Salikoko Mufwene (Chicago)
Pieter Muysken (Leiden)
Peter Mühlhäusler (Adelaide)
Pieter Seuren (Nijmegen)
Norval Smith (Amsterdam)

Volume 21

John McWhorter (ed.)

Language Change and Language Contact in Pidgins and Creoles

LANGUAGE CHANGE AND LANGUAGE CONTACT IN PIDGINS AND CREOLES

Edited by

JOHN MCWHORTER

University of California, Berkeley

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Language change and language contact in pidgins and creoles / edited by John McWhorter.
p. cm. -- (Creole language library, ISSN 0920-9026 ; v. 21)

Fifteen papers selected from those presented at three consecutive meetings of the Society for Pidgin and Creole Linguistics, held Jan. 1996 in San Diego, Jan. 1997 in Chicago, and June 1997 in London.

Includes examples in various creole and pidgin languages

Includes bibliographical references and index.

1. Pidgin languages--Congresses. 2. Creole dialects--Congresses. 3. Languages in contact--Congresses. 4. Linguistic change--Congresses. I. McWhorter, John H. II. Society for Pidgin and Creole Linguistics. III. Series.

PM7801.L36 1999

417'.22--dc21

99-0484886

ISBN 90 272 5243 2 (Eur.) / 1 55619 668 7 (US) (alk. paper)

CIP

© 2000 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. • P.O.Box 75577 • 1070 AN Amsterdam • The Netherlands
John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

Contents

Preface	vii
Verb Movement in four Creole Languages: A Comparative Analysis <i>Marlyse Baptista</i>	1
Notes on Componential Diffusion in the Genesis of the Kabuverdianu Cluster <i>Angela Bartens</i>	35
“High” Kwéyòl: The Emergence of a Formal Creole Register in St. Lucia <i>Paul B. Garrett</i>	63
From Latin to Early Romance: A Case of Partial Creolization? <i>Stéphane Goyette</i>	103
The Creole Verb: A Comparative Study of Stativity and Time Reference <i>John Holm et al.</i>	133
Are Creole Languages “Perfect” Languages? <i>Alain Kihm</i>	163
The Origin of the Syntax and Semantics of Property Items in the Surinamese Plantation Creole <i>Bettina Migge</i>	201
Variable Concord in Portuguese: The Situation in Brazil and Portugal <i>Anthony J. Naro and Maria Marta Pereira Scherre</i>	235
Nativization and the genesis of Hawaiian Creole <i>Sarah Julianne Roberts</i>	257
The Status of Sango in Fact and Fiction. On the one-hundredth anniversary of its conception <i>William J. Samarin</i>	301

Optimality Theory, the Minimal-Word Constraint, and the Historical Sequencing of Substrate Influence in Pidgin/Creole Genesis	335
<i>John Victor Singler</i>	
The Story of <i>kom</i> in Nigerian Pidgin English	353
<i>Sali A. Tagliamonte</i>	
Tense and Aspect in Sranan and the Creole Prototype	383
<i>Donald Winford</i>	
Chaos and Creoles. Towards a New Paradigm?	443
<i>George Lang</i>	
Wh-words and Question Formation in Pidgin/Creole Languages	459
<i>J. Clancy Clements and Ahmar Mahboob</i>	
Index	499

Preface

This anthology collects fifteen papers out of those presented at three consecutive meetings of the Society for Pidgin and Creole Linguistics; namely, those in San Diego in January 1996, Chicago in January 1997, and London in June 1997.

Obviously, choosing these fifteen papers out of the roughly one hundred and twenty-five presented at the three conferences was not an easy task, as among the papers were several dozen fine descriptive studies, historical reconstructions, analyses couched in theoretical phonology, semantics, and syntax, sociolinguistic analyses, and educational treatises.

Wary of allowing my personal interests to bias the selection process, I decided that the guiding factor in gathering the papers for this volume would be originality. I have sought papers which, regardless of their topic, compellingly approach their subject in a novel way, bring to light hitherto under-covered material, or successfully bolster an unconventional or minority case.

The result has been a volume with papers comprising Minimalism, variable rule analysis, Indo-European historical linguistics, substratism, superstratism, the Language Bioprogram Hypothesis, Optimality Theory, language shift, tense-mood-aspect particles, Sango, Indo-Portuguese, even Chaos Theory, and more — united in their aim to keep debate in creole studies moving ever forward in new directions.

I would like to extend sincere thanks to the outside reviewers who so kindly lent their time to commenting on the first drafts of these papers. I offer this volume as a token of appreciation to all of the people in creole studies who have assisted and supported me in my academic endeavors over the past ten years.

John McWhorter
Berkeley, California, July 1999

Verb Movement in four Creole Languages: A Comparative Analysis

Marlyse Baptista

1. Introduction

The goal of this paper is threefold. The first goal is to present a comprehensive syntax of the Capeverdean Creole verb, focusing in particular on the ordering of verbal elements with regard to negation, adverbs, and floating quantifiers. The second is to show how modern syntactic theory can help account for certain descriptive puzzles in Capeverdean Creole syntax; the third is to illustrate how evidence from the verbal syntax of Capeverdean, Haitian, Guinea-Bissau and Louisiana Creoles will necessitate some revisions of current theory of verb movement.

I will focus on the commonly held assumption that there is a strong correlation between verbal morphology and verb movement, or more precisely, that morphologically “rich” subject-verb agreement is responsible for V-raising. Linguists have observed that the poorer the morphology, the more rigid the word order, and it is well known that the inflectional morphology of Creole languages tends to be much simplified in comparison to their lexifier languages (i.e., French, Portuguese, Spanish, or Dutch). Hence, one would not expect creole languages to display verb movement. Yet I will show that contrary to the predictions of various V-raising analyses, some creoles like Capeverdean show evidence of verb movement despite their minimal verbal morphology and lack of overt subject-verb agreement. In this respect, the investigation of the morphology and syntax of creole languages is a particularly interesting topic.

The first part of the paper provides a brief overview of the theory of verb movement, introducing the standard diagnostics for verb movement. In the second section, I argue that there is verb movement in Capeverdean Creole in

spite of minimal verbal morphology and lack of subject-verb agreement. In the third and last part, I compare Capeverdean to other creoles and show the implications of these findings for the theory of verb movement, suggesting revisions of current assumptions and possible new directions for research.

2. The Theory of Verb Movement: Background Assumptions

In recent literature on verb movement, it is common to use the relative position of verbal forms on the one hand and negation, adverbs, and floating quantifiers on the other in attempts to detect verb movement. The arguments for movement go as follows: If there is an element, as in the French examples below (negation in [1a], adverbs in [2a], or floating quantifiers in [3a]), that *precedes* the main verb when it is nonfinite (e.g., the participle in auxiliary constructions), then the verb is assumed not to have moved. If these same elements *follow* the main verb in core (non-compound) tenses, as illustrated by (1b), (2b), and (3b), then the finite verb is assumed to have raised from a “deep” postverbal position to the left of such elements. (3c) shows a quantifier immediately preceding the expression which it modifies.

Negation

- (1) a. *Marie n' a pas lu le livre.* (French)
 Marie NEG has NEG read the book
 ‘Marie has not read the book.’
 b. *Marie ne lit pas le livre.*
 Marie NEG read NEG the book
 ‘Marie does not read the book.’

Adverbs

- (2) a. *Marie n' a jamais vu de fantômes.* (French)
 Marie NEG has never seen any ghosts
 ‘Marie has never seen any ghosts.’
 b. *Marie ne voit jamais de fantômes.*
 Marie NEG sees never any ghosts
 ‘Marie never sees any ghosts.’

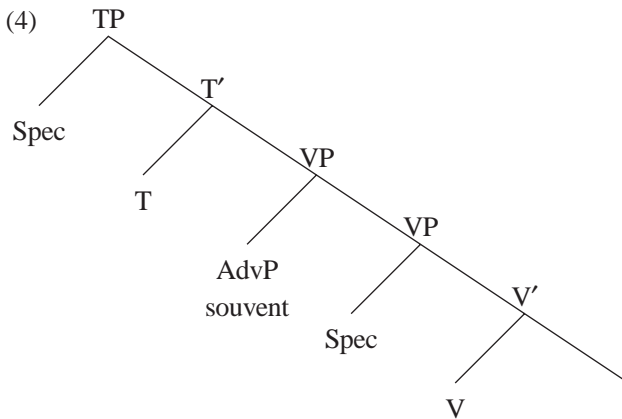
Floating quantifiers

- (3) a. *Les enfants ont tous aimé Jean.* (French)
 the children have all liked Jean
 ‘The children have all liked Jean.’

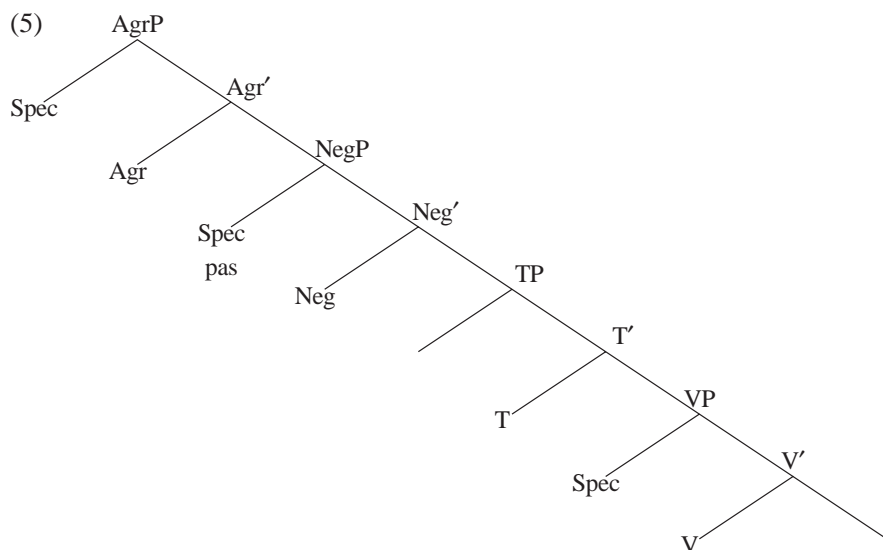
- b. *Les enfants aiment tous Jean.*
 the children love all Jean
 ‘All the children love Jean.’
- c. *Tous les enfants ont aimé Jean.*
 all the children have liked Jean
 ‘All the children have liked Jean.’

These facts are usually analyzed in the following way: The main verb is assumed to be generated inside the VP; and elements that precede it in the contexts which do not condition movement are either left-adjoined to the VP, as in the case of French adverbs under the phrase structure given in (4); or have their own projection higher than the verb (as in the case of *pas* located in Spec-NegP in French, as in [5]); or occur in Spec-VP (the quantifier may be left behind by a subject that is commonly assumed to be base-generated in Spec-VP, as shown in [6]). When it is finite, the main verb is then believed to move to some higher projection, such as I^0 , or T^0 , or AgrS (depending on the framework or the type of clausal structure assumed), in languages that have V-movement.

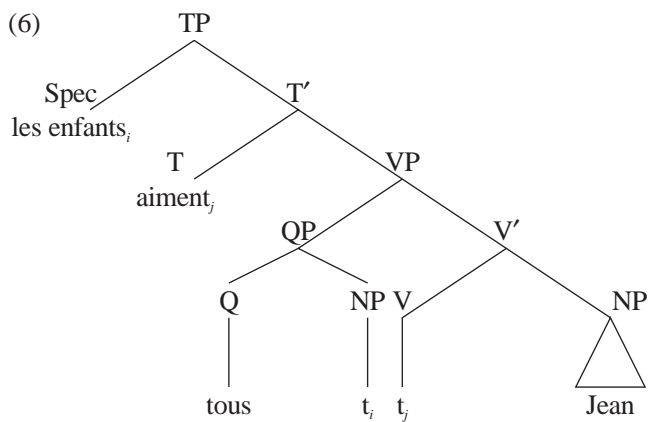
Adverbial left-adjunction:



Neg heading its own projection:



Floating quantifier:



The general theoretical question arises as to why languages such as French (and Icelandic, Old English, Old Norse, Old Swedish / Danish, and Norwegian among others) seem to have V-movement, whereas others, such as English, as illustrated in the following examples, do not:

English:

- (7) a. *John did **not like** the apple pie.* (Negation)
 b. **John **liked not** the apple pie.*
- (8) a. *John **often tells** stories.* (Adverb)
 b. **John **tells often** stories.*¹
- (9) a. *The guests **all saw** Mary crying.* (Floating quantifier)
 b. **The guests **saw all** Mary crying.*

As illustrated by the ungrammaticality of (7b), (8b), and (9b), finite main verbs² in Modern English may occur only in a post-Neg position.

This leads us to consider the competing approaches to this cross-linguistic variation.

3. Review of Approaches to Verb Movement

3.1. *Split IP*

Pollock (1989) suggests a highly articulated structure of IP and adduces some empirical evidence that inflection is split into AgrP and TP. He also assumes the maximal projection NegP. Pollock's basic assumption is that the properties of these new projections determine whether or not a verb can move. Hence, he assumed that long verb movement from V⁰-to-T⁰ to Agr⁰ is allowed in French, due to the "transparency" of its AGR, whereas long verb movement in English is prohibited due to the "opacity" of its AGR. In other words, if AGR is transparent, verb movement is allowed; and if AGR is opaque, movement is not allowed. The concepts of transparency and opacity were reinterpreted by Chomsky (1993) in terms of "strong" and "weak" features.³ When the V-features are strong, the verb must move overtly; when they are weak, the verb moves covertly.

3.2. *Strength of Verbal Inflection*

Vikner (1992, 1995) has consistently emphasized a correlation between the strength of verbal inflectional morphology and the obligatory movement of the finite verb to I⁰ (i.e., to the left of a medial adverbial or negator). In an attempt to articulate a typology of features that trigger verb movement, Vikner (1995) argued that V⁰-to-I⁰ occurs only if all core tenses (meaning noncompound tenses) in a given language are inflected for person.

Rohrbacher (1995) adopts a more restrictive view than Vikner (1995) and argues that V⁰-to-I⁰ movement occurs if and only if 1st and 2nd person are overtly and distinctly marked at least once (that is, for 1st and 2nd person singular or 1st and 2nd person plural) in a given tense. Rohrbacher's most recent proposal (1995: 363) is that languages are not parameterized for V-to-AgrS raising and pro-drop as such, but rather for the presence or the absence of lexical entries for their agreement affixes. Whereas a positive setting of this parameter will trigger V-to-AgrS raising and allow pro-drop, a negative setting will prohibit both. In other words, the trigger for V⁰-to-I⁰ movement lies in the agreement paradigm of a given language.⁴

Following Platzack (1988), Roberts (1993) draws evidence from the history of English and Mainland Scandinavian, among other languages, supporting the view that verb movement is associated with rich verbal morphology. He corroborates Platzack's (1988) observation that the loss of subject-verb agreement in both English and Swedish was contemporaneous with the loss of V⁰-to-I⁰ movement and occurred within the 200-year period between 1500 and 1700 (Platzack 1988: 223).

Consider the following examples from Middle English / Early Modern English, where the occurrence of the order V-*not* (as in [10]), V-*adverb*, and V-*quantifier* has been interpreted as resulting from V⁰-to-I⁰ movement:

Middle English (1100-1500):

- (10) a. *Wepyng and teres counforteth not dissolute laghers.*
 'Weeping and tears did not comfort the careless laughers.'
 (1400-1450: N. Love: *The Myroure of the Blessyd Lyf of Jesu Christ* [Gray 1985: 97])

Early Modern English (1500-1800):

- b. *They were ful soore adredde and wist not what it was.*
 'They were so afraid and did not know what it was.'
 (1438: Anon: *The Gilte Legende* [Gray 1985: 103])

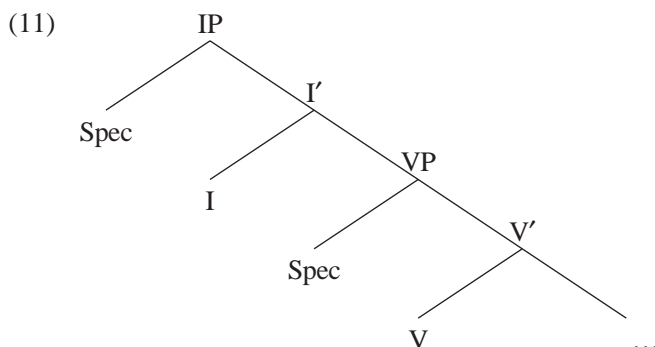
On the basis of such evidence, Roberts concludes that Middle English and Early Modern English in its early stages (up to 1550-1575) required V^0 to move to I^0 in all tensed clauses and lost this type of movement later on. This added further evidence linking “rich” verbal morphology to verb movement.

3.3. Split IP-Parameter

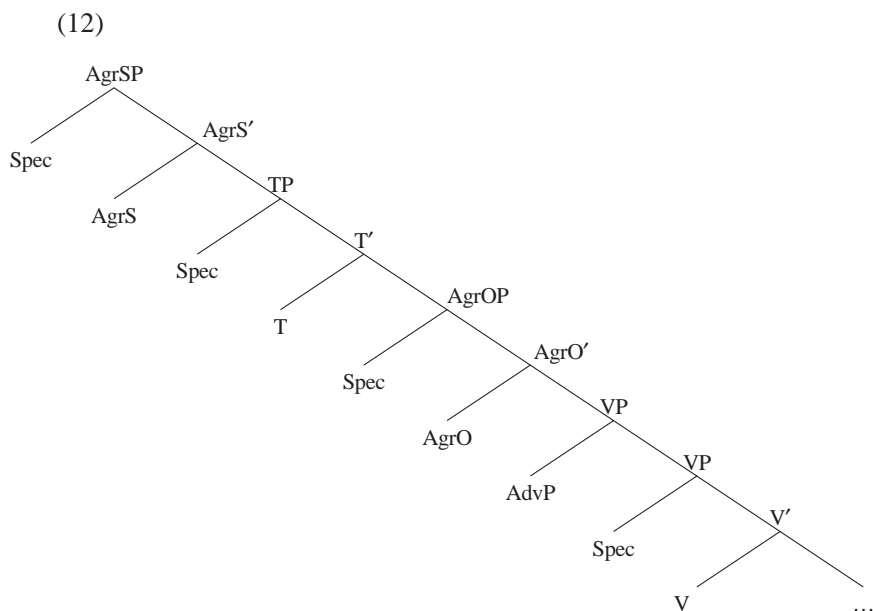
Thráinsson (1996: 267) correctly notes that it is difficult to define the concept of “rich” morphology in such a way that it makes the correct predictions with regard to overt verb movement. Instead of assuming a direct correlation between rich verbal morphology and overt verb movement, Thráinsson assumes a correlation between rich verb morphology and Split IP, under which TP is a syntactic category separate from AgrSP.

His argument was as follows: In the absence of Split IP, the simplified IP structure will emerge and no AgrSP/TP/AgrOP will appear. This would account for the different number of subject and object positions available in languages with rich verbal morphology versus those with impoverished verbal morphology. Indeed, languages with just IP above the VP would have only Spec-IP as an available subject position above VP, whereas languages with Spec-AgrSP and Spec-TP would have two positions (and also a Spec-AgrOP for the shifted object). Bobaljik & Thráinsson (1997) argue that the verb needs to move to T^0 in a Split IP structure but does not raise to I^0 in an unsplit IP structure. They opted for a structural account rather than one that relies on differential feature strength. Consider the following two structures:

Unsplit IP:



Split IP:



They pointed out that in the unsplit IP structure in (11), the VP headed by the V^0 is the complement of I^0 . If an AdvP is adjoined to VP, it does not alter that relationship because it does not create any new projection intervening between VP and I^0 . If one assumes that the complement of a head is in its checking domain, then the verb does not have to move to check features present in I (Tense or other features).⁵ Bobaljik and Thráinsson propose, however, that when other projections intervene between the verb and the head(s) it has to check features with, the verb has to raise to get into a checking relationship with the relevant head. This is illustrated in (12). In that structure, the verb has to check features with T^0 , at least, but AgrOP intervenes, so the verb must raise overtly to T^0 .

In this respect, Thráinsson (1996: 279) questions whether a language lacking inflectional morphology but displaying a Split IP (a TP separate from an AgrSP) could be found.⁶ I show in this paper that although Capeverdean is endowed with minimal inflectional morphology (the single suffix *-ba*), it gives evidence of overt V-raising. Following Thráinsson's line of reasoning, it could be argued then that Capeverdean Creole has a Split IP (we will return to this issue subsequently).

In the next section, we turn to the diagnostics for V-raising in Capeverdean, and examine the distribution of verbs and TMA markers *vis-à-vis* the negator *ka*.

4. Position of Verbs and TMA Markers

4.1. The Position of Verbs and TMA Markers in relation to *ka*

When expressing sentential negation, Capeverdean *ka* precedes not only the main verb, as shown in (13), but also the sequence of TMA markers. In other words, *ka* must be preverbal and never allows *ta*, as in (14), *sta*, as in (15), or the combination *sta ta*, as in (16)-(17) to precede it.

- (13) *João ka kume katxupa.*
 João NEG ate katxupa
 ‘João did not eat any katxupa.’
- (14) a. *João ka ta kume karni.*
 João NEG TMA eat meat
 ‘João does not eat meat.’
 b. **João ta ka kume karni.*
 João TMA NEG eat meat
- (15) a. *João ka sta kume karni.*
 João NEG sta eat meat
 ‘João is not eating meat.’
 b. **João sta ka kume karni.*
 João TMA NEG eat meat
- (16) a. *João ka sta ta kume karni.*
 João NEG TMA TMA eat meat
 ‘João is not eating meat.’
 b. **João sta ka ta kume karni.*
 João TMA NEG TMA eat meat
- (17) a. **João sta ta ka kume karni.*
 João TMA TMA NEG eat meat
 b. **João sta ta kume ka karni.*
 João TMA TMA eat NEG meat

The generalization that we can derive from these data is that whether *ta* and *sta* occur alone, as in (14a) and (15a), or in combination as in (16a), these markers must follow negation.

There is, however, one interesting exception to this generalization: whereas all verbs follow negation, the copula *e* generally appears in a pre-Neg position.⁷ This sole exception is worthy of further analysis and is the topic of the next subsection.

4.2. *The Position of e in relation to ka*

The morpheme *e* is pre-Neg (in most dialects)⁸ and allows the negative morpheme to immediately precede adjectival predicates as in (18) and nominal predicates as in (19):

- (18) a. *João e ka temozu.*
 João e NEG temozu
 ‘João is not stubborn.’
 b. **João ka e temozu.*
 João NEG e stubborn.
- (19) a. *João e ka nha pai.*
 João e NEG my father
 ‘João is not my father.’
 b. **João ka e nha pai.*
 João NEG e my father

Such constructions abound in Capeverdean literature, as attested by examples such as (20):

- (20) *Kauberdi e ka mas e ka ménus ki un txabasku*
 Cape Verde is NEG more is NEG less than a piece
 *di Goltarpu.*⁹
 of Goltarpu
 ‘Cape Verde is no more, no less than a colony of Goltarpu.’
 (Veiga 1987:14)

This is somewhat similar to the exceptional, pre-Neg position of the (inflected) English copula, but contrary to the English case, the past tense counterpart of *e*, *era*, is always post-Neg, whether in adjectival or nominal

predicates, as illustrated by the examples in (21) and (22):

- (21) a. *João ka era temozu.*
 João NEG was stubborn
 ‘João was not stubborn.’
 b. **João era ka temozu.*
 João was NEG stubborn
- (22) a. *João ka era nha pai.*
 João NEG was my father
 ‘João was not my father.’
 b. **João era ka nha pai.*
 João was NEG my father

The same situation arises for the future tense counterpart of *e*, *ta ser*. *Ta ser* must assume a post-Neg position, as is shown by the ungrammaticality of (23b) and (23c):

- (23) a. *João ka ta ser profesor.*
 João NEG TMA be professor
 ‘João will not be a professor.’
 b. **João ta ser ka profesor.*
 João TMA be NEG professor
 c. **João ta ka ser profesor.*
 João TMA NEG be professor

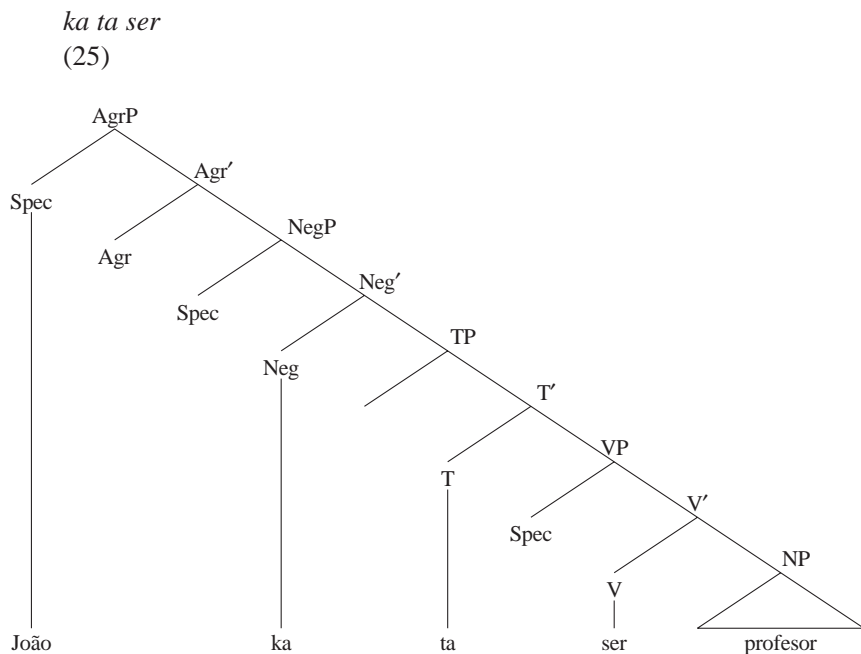
This state of affairs is summarized by the templates in (24):

- (24) a. *ka ser* + Present = *e ka*
 b. *ka ser* + Past = *ka era*
 c. *ka ser* + Future = *ka ta ser*

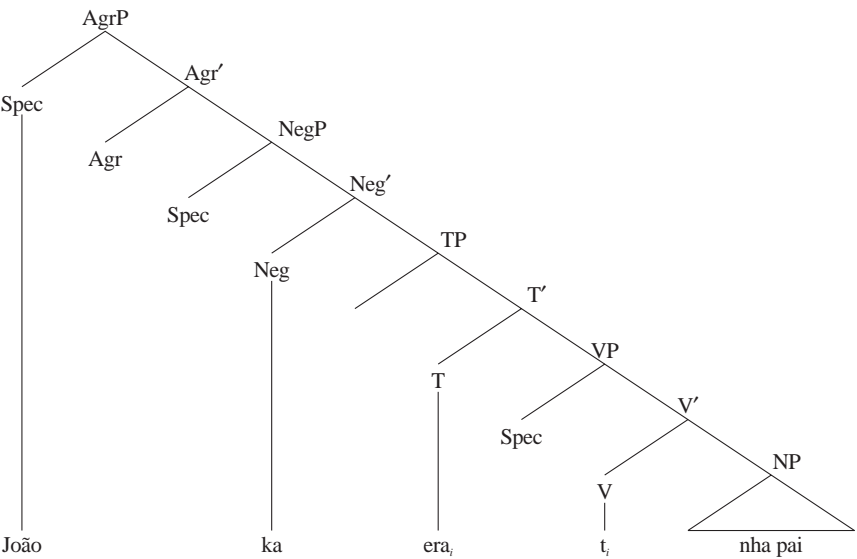
Chomsky (1993) argued that auxiliaries raise in English due to their semantic vacuity; lacking semantically relevant features, they are not visible to LF rules. Such arguments cannot hold for Capeverdean Creole. Indeed, the movement of *e* to a pre-Neg position cannot be explained by assuming that its raising is triggered by its auxiliary status, given that auxiliaries such as *ta* and *sta* can never be found in a pre-Neg position, as shown in (14b) and (15b). We may infer from this that *e* is not an auxiliary but that its morpho-phonological lightness may be triggering its raising.¹⁰

The trees in (25), (26) and (27) illustrate the assumed position of, respec-

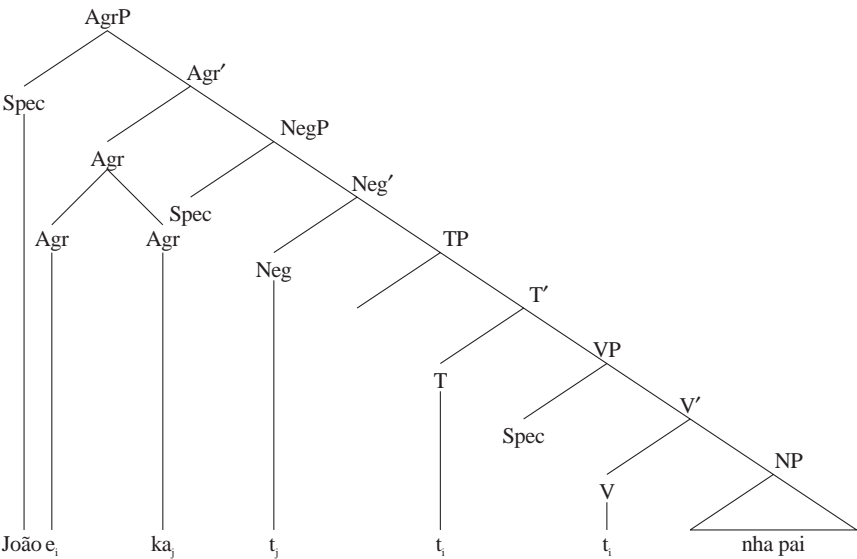
tively, the future, past and present tense forms of *ser* ‘be’. For ease of exposition, we assume that *ka* heads the projection NegP; the tree in (25), representing the future tense, shows that *ser* in the infinitival form does not have any reason to move, as it has no inflection to pick up in T⁰. Hence, the main verb remains *in situ*. In contrast, the tree in (26), representing the past tense, shows that the verb has moved from V⁰ to T⁰, presumably to pick up a past tense feature. The tree in (27), representing the present tense, shows that the copula may have moved overtly from V⁰-to-T⁰ to Agr⁰ to a pre-Neg position. First, *e* moves to Neg-head and adjoins to *ka*; and then the complex [*e ka*] moves to Agr⁰. The fact that *ka* may raise in some constructions is reminiscent of the French negative particle *ne*, which raises to Agr⁰, and the Italian morpheme *non*, which also raises to Agr⁰, as described in Belletti (1990).



ka era
(26)



e ka
(27)



In summary, in this subsection, we have observed that verbs and TMA markers are always in a post-Neg position. Verbs raise covertly from V^0 -to- T^0 and remain in a post-Neg position, whereas the copula *e* raises overtly from V^0 -to- T^0 to Agr^0 , landing in a pre-Neg position. On this issue, I give evidence in the next subsections that verbs can move overtly from V^0 -to- T^0 past a certain class of adverbials and floating quantifiers.

4.3. Capeverdean Verb Position With Regard to Adverbs

As discussed in the section on the theory of verb movement, it is generally assumed that whether the finite verb is in V^0 or not can be determined from its position relative to a sentence-medial adverbial (i.e., an adverbial that follows the subject but precedes the complement of the verb). The medial adverbial is assumed to left-adjoin to VP. This means that if the verb precedes the adverbial, it has left VP; whereas if the verb follows the adverbial, it must still be in V^0 .

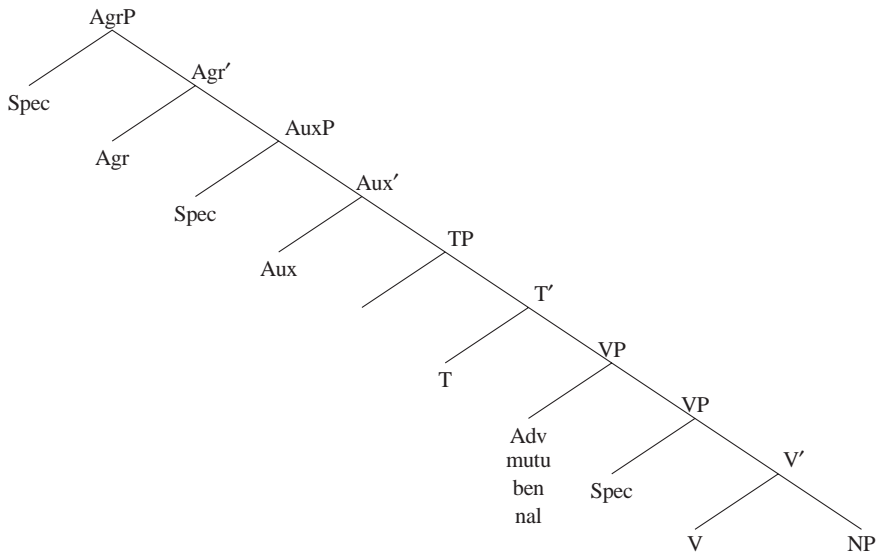
Capeverdean has a class of adverbs that occur preferably in a postverbal position, and possibly sentence-finally (in some dialects), as illustrated by (28) – (30). The following sets of sentences show that this class of adverbs (*mutu* ‘too much’/‘a lot’; *ben* ‘well’, *mal* ‘badly’) occur postverbally, as we see in the (a) examples; the ungrammaticality of the (b) examples shows that these adverbs cannot occur between the subject and the verb. The (c) examples show that they cannot occur sentence-initially. The question mark in the (d) examples expresses that sentence-final occurrence of these adverbs is possible in some dialects and marginal in others.

- (28) a. *João ta ama mutu Eliza.*
 João TMA love too much Eliza
 ‘João loves Eliza too much.’
 b. **João mutu ta ama Eliza.*
 João too much TMA love Eliza
 c. **Mutu João ta ama Eliza.*
 too much João TMA love Eliza
 d. ?*João ta ama Eliza mutu.*
 João TMA love Eliza too much
- (29) a. *João xina ben se lison.*
 João learned well his lesson
 ‘João learned his lesson well.’

- b. **João ben xina se lison.*
 João well learned his lesson
- c. **Ben João xina se lison.*
 well João learned his lesson
- d. ?*João xina se lison ben.*
 João learned his lesson well
- (30) a. *João xina mal se lison.*
 João learned badly his lesson
 'João learned his lesson badly.'
- b. **João mal xina se lison.*
 João badly learned his lesson
- c. **Mal João xina se lison.*
 badly João learned his lesson
- d. ?*João xina se lison mal.*
 João learned his lesson badly

We assume that members of this class of adverb are generated as left-adjuncts to VP, as illustrated in (31):

(31)



Assuming that these adverbs are base-generated left-adjoined to VP and that adverbs are not subject to movement, the preadverbial position of the verb would indicate that the verb has moved to T⁰. The only other way to derive this word order would be to say that the object was shifted to the right, as it occurs in heavy NP shift constructions. However, an NP like *Eliza* in (28a) is not the kind of heavy NP or new-informational NP one would expect to move to the right, so this does not seem to be an option. Consequently, this type of adverb provides us with a crucial test showing that the verb has moved to T⁰.

There is, however, an important observation that should be made regarding the placement of such adverbs with respect to Capeverdean auxiliaries. Consider (32), the analogue to (29) except for the presence of the auxiliaries *sta* and *ta* and the anterior marker *-ba*.

- (32) a. *João staba ta xina ben se lison.*
 João TMA+ba TMA learn well his lesson
 ‘João was learning his lesson well.’
 b. **João ben staba ta xina se lison.*
 João well TMA+ba TMA learn his lesson
 c. **Ben João staba ta xina se lison.*
 well João TMA+ba TMA learn his lesson
 d. ?*João staba ta xina se lison ben.*
 João TMA+ba TMA learn his lesson well

(32) shows that whether in a compound or noncompound tense, the adverbial must be immediately postverbal, as illustrated by the ungrammaticality of (32b-c) and the marginality of (32d).

Recall that in French, the adverbial appears in compound tenses (involving auxiliaries) between the finite auxiliary and the past participle, as in (2a). It was observed that if the relevant functional head position (I⁰ or T⁰, or Agr⁰) is filled with an auxiliary element, then the main verb remains *in situ*; but if no such auxiliary element is present, then the main verb moves to that functional head position in core (noncompound) tenses. In other words, verb movement in French (and other languages such as Icelandic) affects main verbs only when no auxiliaries are present. The reason for this is that auxiliary verbs in French are inflected for tense and agreement, just like ordinary verbs, so it is these verbs that check the relevant features when they are present. Hence, it is important to emphasize that contrary to the case in French, the Capeverdean adverbial does not behave differently in compound and noncompound tenses,

as shown by example (29) where auxiliaries are absent and example (32) where auxiliaries are present.

In relation to this, I argue in Baptista (1997) that the element *ta* is a nonverbal auxiliary (it cannot carry *-ba*). Hence, there is no reason to expect that it will interfere with the raising needs of the main verb. Second, I assumed a biclausal structure which supports the assumption that the element *sta* takes a clausal complement which can take its own *ta* and *-ba* markers, as shown in the tree in (40). Thus it is possible that checking in the complement of *sta* would still need to be done by the main verb. These assumptions explain why, contrary to the case in French, the presence of auxiliary elements like *ta* or *sta* does not have any effect on the raising possibilities and raising needs of the main verb.

Let us now turn to some evidence from floating quantifiers which provide another diagnostic for verb movement.

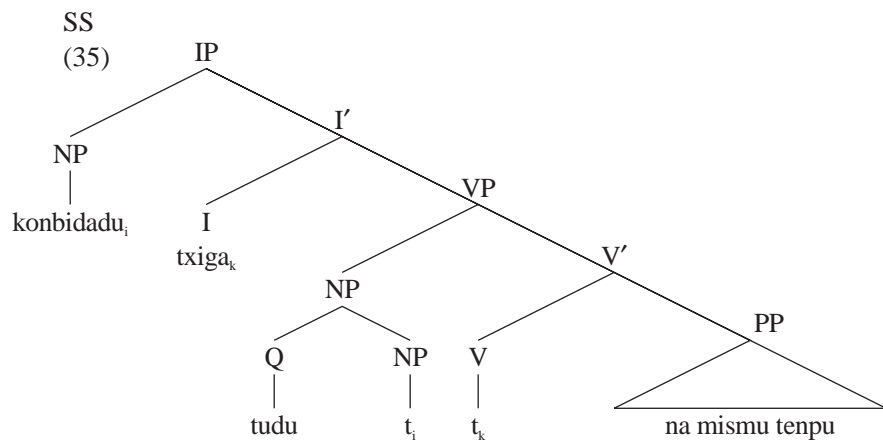
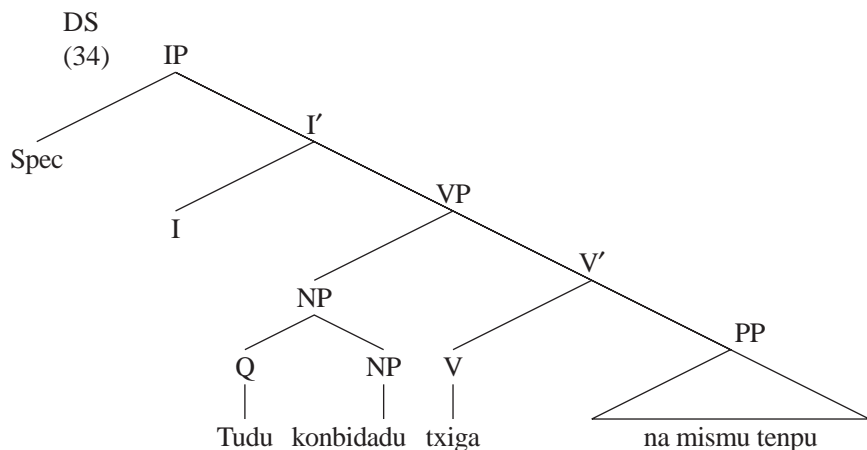
4.4. Quantifier Float in Capeverdean

Capeverdean displays the same type of quantifier float as French; hence, a floating quantifier may be preverbal, as in (33a), or postverbal, as in (33b). (33b) provides us with crucial evidence that the verb has moved to T^0 , given that the verb precedes the floating quantifier which has remained in its DS position. This yields the tree representations in (34) and (35).

- (33) a. ***Tudu konbidadu txiga na mismu tenpu.***
 all guests arrived in same time
 ‘All the guests arrived at the same time.’
 b. *Konbidadu txiga tudu na mismu tenpu.*
 guests arrived all in same time
 ‘All the guests arrived at the same time.’
 c. *Konbidadu tudu txiga na mismu tenpu.*
 guests all arrived in same time

The assimilation of the behavior of a floating quantifier to that of VP-initial adverbs follows from the fact that they have the same DS (deep structure) location. In Sportiche’s (1988) account, the NP *konbidadu* must move at SS to a position where it can be case-marked. That is why at SS (surface structure) it fills the typical subject position that has been identified with Spec-IP (or Spec-AgrP). As the subject moves, the modifying quantifier

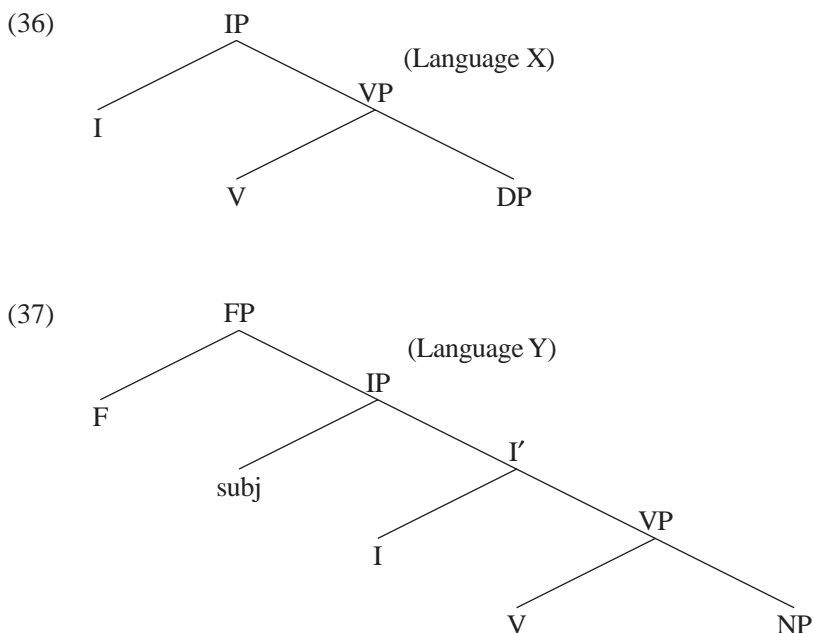
may remain in place, as illustrated in (35). The tree in (35) crucially shows that the verb has moved to I⁰ (or to T⁰ in our framework)¹¹, past the quantifier. In this respect, floating quantifiers, like VP-adjoined adverbs, provide clear evidence of V-raising in Capeverdean Creole.



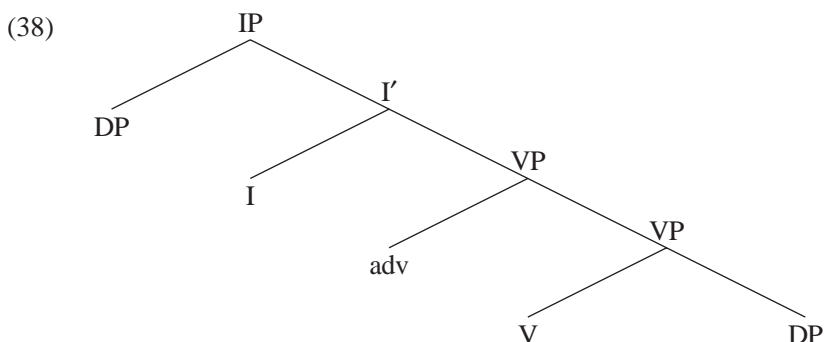
5. Comparative Creole Observations

5.1. A New Lead?

We will first examine the approach to V-raising adopted by Bobaljik (1995) and Bobaljik and Thráinsson (1997). Consider the two trees in (36) and (37), from Bobaljik (1995: 268-269):



Regarding Language X, represented by the tree in (36), Bobaljik's main proposal is that if a language has only one head with V-features, in this case I^0 with VP as its sister, then the verb remains in VP. This is because Infl and VP are in a local relationship, and the verb therefore has no motivation to raise for the purpose of feature-checking. (In other words, V^0 and Infl are in a checking relationship when they are adjacent.) Bobaljik (1995:275, fn. 17) added that given the adjunction structure usually represented as in (38), adverbs do not disrupt the adjacency relationship; only specifiers and heads do.



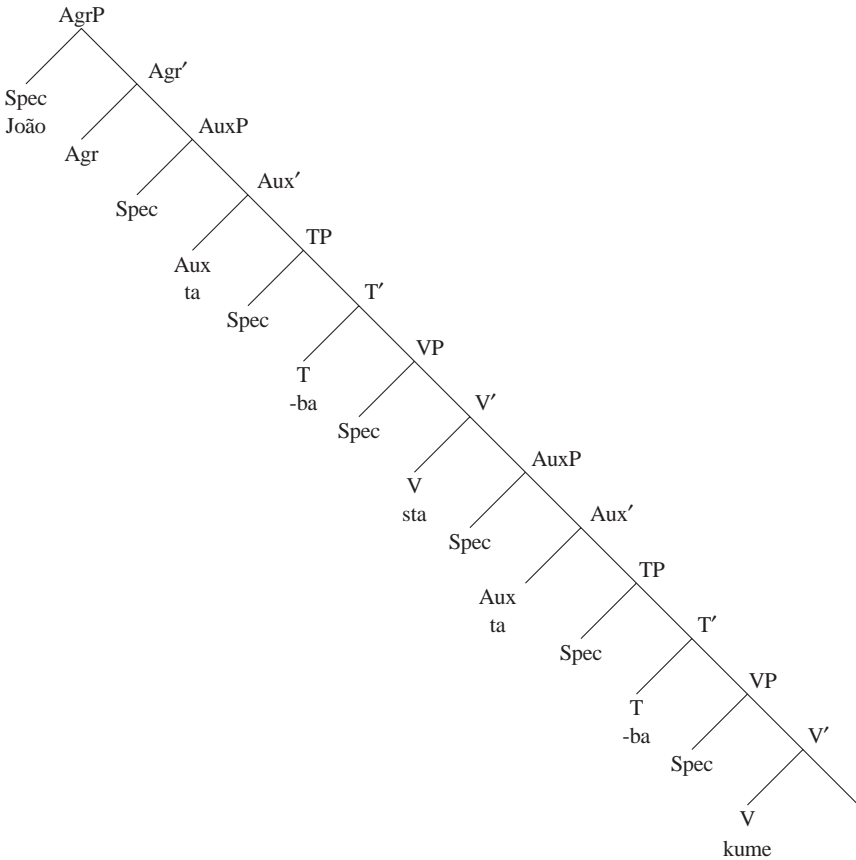
If, on the other hand, Language Y (in tree (37)) has two or more heads with V-features above the VP, then checking is not satisfied without movement; the verb will have to raise to the highest of these heads, the head F in (37).

Bobaljik and Thráinsson (1997) brings further modification to this proposal. Under this theory, they assume that there is no reason for V^0 to move all the way to F in (37) rather than just to I^0 . Once it is in I, it heads the complement of F, and if the Head-Complement relation is a checking relation, as we assume, then the V could check features against F even if it did not move any further than I^0 . Their line of reasoning reflects a move away from the correlation between V-raising and morphology.

Let us now draw a parallel with the Capeverdean case and propose a tentative structure. I assume a biclausal architecture for Capeverdean. Thus the sentence in (39) is represented by the tree in (40):

- (39) *João ta staba ta kumeba.*
 João TMA sta+ba TMA eat+ba
 'João would have been eating.'

(40)



This leads us to the next question: Why does the Capeverdean verb need to raise overtly to T^0 ? We have seen ample evidence for this movement in the previous section; in this respect, it is of interest to compare Capeverdean to Haitian, as we can derive interesting theoretical hypotheses from such a comparison.

5.2. Haitian Creole

The syntax of the Haitian verb is described thoroughly in DeGraff (1997). Haitian has no overt subject-verb agreement (DeGraff, 1996: 11). Haitian has

TMA markers which are all preverbal and, crucially, the language has no verbal suffixes. In the presence of VP-internal adverbials, the verb always remains in V⁰ (ibid. 17).

Consider the Haitian sentence in (41). The Haitian verb cannot raise past VP-adjoined adverbs, as shown by the ungrammaticality of (41b).

- (41) a. *Bouki te ap mal manje.* (Haitian)

Bouki TMA TMA badly eat

‘Bouki was eating badly.’

- b. **Bouki te ap manje mal.*

Bouki TMA TMA eat badly

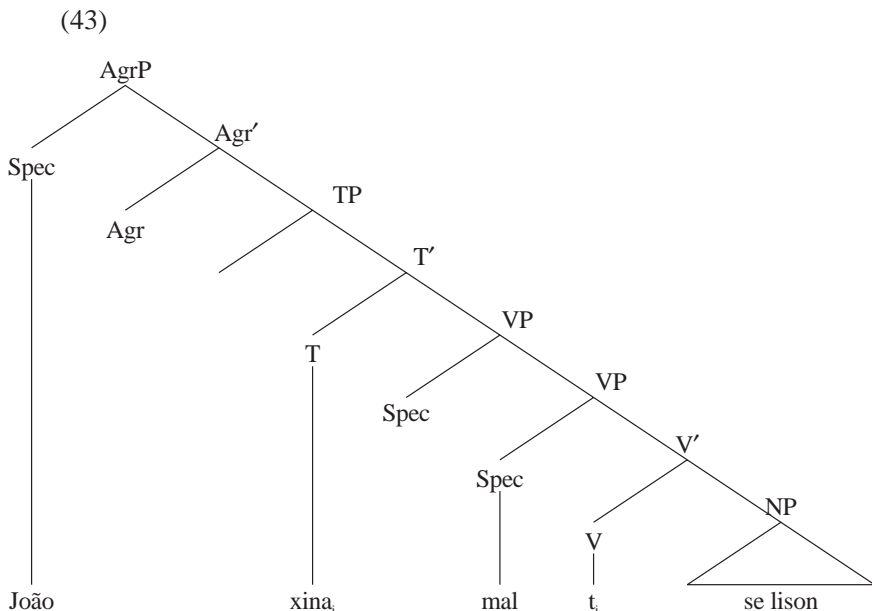
The difference in verbal behavior between Capeverdean and Haitian would at first seem to be due to the suffix *-ba* in Capeverdean, for which no equivalent exists in Haitian. This would lead us to view *-ba* as the trigger for the syntactic movement of the verb to T⁰.

This is not an uncontroversial issue. Koopman (1984) and Bobaljik (1995) noted that the property [+affix] as a trigger for syntactic movement (in this case V-movement) is a problematic assumption, as movement occurs regardless of the morphophonological content of the affix. More precisely, raising occurs even when the supposed affix trigger is not phonological. That is indeed the case in Capeverdean, where the bare stem of nonstative verbs get a simple past tense and do raise in spite of the lack of suffixation. This is illustrated in (43).

- (42) *João xina mal se lison.*¹² (Capeverdean)

João learned badly his lesson

‘João learned his lesson badly.’



We come back to this problem in a section below, where we consider the behavior of verbs in Louisiana Creole.

This leads us to explore Bobaljik and Thráinsson's alternative approach. Let us first discuss the Capeverdean case. If one assumes the structure suggested in (43), with a TP and an AgrP as different functional projections, and if one assumes furthermore that the verb in the VP needs to check some features with the Agr-head, then it will have to move to T^0 to do so. The fact that the anterior marker *-ba* behaves differently from *ta* may suffice as evidence to the language learner that different elements, and hence plausibly different functional heads in the syntax, are involved. In Haitian on the other hand, in the absence of a suffix, the learner deduces that the verb does not have any raising needs.

Following Bobaljik & Thráinsson (1997), then, this crucial difference between Capeverdean and Haitian is similar to the difference between French or Icelandic on the one hand, and English and Mainland Scandinavian on the other. In Capeverdean as well as French and Icelandic, we have evidence for two kinds of functional projections above the VP, namely, an AgrP and a TP. Thus the verb has to raise out of the VP in order to check features against the

- (45a) shows that the position preceding the adverb is the only position in which definite subjects may surface, while (45b) reveals that the position

following the adverb is only available for indefinite subjects. This leads Jonas & Bobaljik to argue that indefinite subjects are not internal to the VP at S-structure, and that indefinite subjects are in the specifier of an intermediate functional projection Spec-TP. In addition, Icelandic displays structures of the type “The troll my hat ate”, where the object “my hat” has undergone overt object shift and been raised to Spec-AgrO. Neither sentences of this kind nor ones such as (44) or (45) are found in Capeverdean Creole, and thus there is no reason to assume empirically that Capeverdean would be endowed with a Spec-TP and a Spec-AgrOP.

The last step in this analysis is to look at V-raising in Capeverdean from a developmental perspective. In the same way that the loss of V^0 -to-Agr⁰ movement in English and Mainland Scandinavian accompanied the loss of verbal inflection (Roberts 1993), Capeverdean may possibly have gained short V-movement after acquiring a verbal suffix, an unusual trait among creole languages. The scarcity of historical texts in this creole will make it challenging to ascertain whether *-ba* was initially an independent morpheme in Capeverdean, but uncovering such evidence would be useful in confirming the link between the appearance of suffixation and verb movement.

In contrast, *ba* is found frequently as an unbound morpheme in Guinea-Bissau Creole (Kihm 1994), a Creole closely related to Capeverdean. In the next section, we consider the case of Guinea-Bissau and examine the morpheme *ba* and show how the verb behaves with regard to VP-adverbials and floating quantifiers.

5.3. Guinea-Bissau Creole

As described in Kihm (1994), Guinea-Bissau Creole has three TMA markers, *ba*, *na* and *ta*. The behavior of *ba* differs from that of *na* and *ta* in three ways: First, it follows the item it modifies. Secondly, it does not require adjacency to this item. Thirdly, this item may be a noun predicate, an adjectival predicate as in (46) and (47), a verb as in (48) and even a nonpredicate constituent, as shown in (49):

- (46) *N ta kontenti ba na kil tenpu.* (Guinea-Bissau Creole)
 I TMA happy [+Past] in that time
 ‘I was happy in that time.’ (Kihm 1994: 102)

- (47) *I un prosesu difisil ba.*
 it a process difficult [+Past]
 'It was a difficult process.' (Kihm 1994: 108)
- (48) a. *N kontu u ba kuma nya pirkitu karu de.*
 I tell you Past that my parrot expensive de
 'I had told you though that my parrot is expensive.' (Kihm 1994: 99)
- b. *E gaša li i ten un kaw ku nciga di oja l*
 this girl here it have a place that I arrive of see her
ba nel.
 Past in.it
 'This girl, I already saw her some place.' (Kihm 1994: 99)
- (49) *I ka el ba.*
 it NEG s/he/it Past
 'It was not him/her/it.' (Kihm 1994: 99)

There are, however, instances where *ba* occurs immediately after the verb, displaying a distribution similar to that of Capeverdean, as shown in (50):

- (50) *A! N diskisi ba!*
 ah I forget Past
 'Ah I had forgotten!' (Kihm 1994: 99)

Kihm (1994: 101) notes that these occurrences of *ba* represent more marginal, decreolized varieties, in which *ba* is more used as an imperfect ending, parallel to Portuguese *-va*. He analyzes *ba* not as a verb or adverb, but rather as a [+N, +V] item with an added feature [+ADV] (ibid. 105).

The above description of *ba* shows that this morpheme does not behave as a canonical verbal affix. Its nonverbal affix status should allow us to make two correlated predictions — 1) *ba* is not in T, and hence cannot trigger V-raising, and that 2) we should therefore not expect V-raising in Guinea-Bissau Creole. Closer examination of the facts shows that these predictions are only partially borne out. Let us examine the position of the Guinea-Bissau verb *vis-à-vis* negation, VP-adverbials and floating quantifiers.

With regard to negation, Guinea-Bissau *ka* behaves just like Capeverdean *ka*. When it modifies a verb, it always immediately precedes it, as shown in (51):

- (51) *Ze ka riba inda.*
 Ze NEG return yet
 ‘Ze has not returned yet.’ (Kihm 1994: 42)

In this respect, just as with ordinary verbs in Capeverdean, *ka* cannot be used as a diagnosis for V-raising. However, a contrast between the two creoles arises with respect to VP-adverbials. Indeed, as illustrated by the examples in (52) and (53), the VP-adverbials *ciw*, ‘a lot’, and *diritu*, ‘well’, cannot occur in a postverbal position, as shown in (52b) and (53b) respectively (Kihm, personal communication)¹³:

- (52) a. *Jon ta kiri Eliza ciw.*
 Jon TMA like Eliza a lot
 ‘Jon like Eliza a lot.’
 b. **Jon ta kiri ciw Eliza.*
 Jon TMA likes a lot Eliza¹⁴
- (53) a. *Jon prindi si lison diritu.*
 Jon learned his lesson well
 ‘Jon learned his lesson well.’
 b. **Jon prindi diritu si lison.*
 Jon learned well his lesson¹⁵

Hence, as we predicted, the Guinea-Bissau verb remains *in situ* and does not move past VP-internal adverbials. Our predictions, however, are not borne out with regard to floating quantifiers. Indeed, the Guinea-Bissau Creole quantifier *tudu* can be stranded and the verb can raise past it, just as in Capeverdean. This is shown in (54):

- (54) a. *Konbidadu ciga tudu na mismu tenpu.*
 guests arrived all at same time
 ‘The guests arrived all at the same time.’
 b. *Konbidadu tudu ciga na mismu tenpu.*
 guests all arrived at same time
 ‘The guests arrived all at the same time.’¹⁶

Thus the data in (54) challenges the predictions we made that the Guinea-Bissau Creole verb would not move due to having no inflection to pick up or any feature to check in T. These data show us that there may be more to the general verb movement picture than we have assumed so far.

5.4. Louisiana Creole

As described in Rottet (1992), there is in mesolectal Louisiana Creole a morphosyntactic alternation between full and truncated verb stems which is absent in the basilectal varieties. More precisely, in mesolectal Louisiana Creole, the alternation is \emptyset versus *-e*. Hence, a verb like *mōzhe* “to eat”, can alternate between the full stem *mōzhe* and the truncated stem *mōzh*.¹⁷ Basilectal Louisiana Creole is only endowed with the full stems.

Rottet (1992), drawing most of his data from Neumann (1985, 1987), notes that only the short verb stems undergo verb movement, whereas the full verb stems do not. For instance, in negative constructions, Rottet notes that the long stem form does not move, hence remains in a post-Neg position, whereas the short stem form moves and appears in a pre-Neg position. This is illustrated in (55a) and (55b) respectively:

- (55) a. *Na lōtō mo pa mōzhe gratō.*
 in long time I NEG eat cracklin’
 ‘I haven’t eaten cracklin’ for a long time.’
 b. *Mo mōzh pa gratō.*
 I eat NEG cracklin’
 ‘I don’t eat cracklin’.’ (Neumann 1985:321, cited in Rottet 1992: 277)

Short and long verb stems also show a discrepancy with regard to VP adverbs, such as *zhame*, “never”. Such adverbs must precede the long verb stem, as in (56), whereas they can occur before or after the short verb stem, as illustrated in (57):

- (56) a. *Mo (te, se, sa, ...) zhame zhōngle ōho sa.*
 I (ANT, IRR, FUT) never think about that
 ‘I never thought/would think/will have thought about that.’
 b. *Mo(pa) zhame (te,...) zhōgle ōho sa.*
 I (NEG) never (ANT,...) think about that
 ‘I never thought about that.’
 (Neumann 1985: 330, cited by Rottet 1992: 267)
- (57) a. *Mo zhame marsh ni-pje deor.*
 I never walk barefoot outside
 ‘I never walk barefoot outside.’

- b. *Mo marsh (pa) zhame ni-pje deor.*
 I walk (NEG) never barefoot outside
 'I never walk barefoot outside.'
 (Neumann 1985: 330, cited by Rottet 1992: 267)

This leads us to a second observation: The occurrence of verb stems in a pre-Neg position are an innovation in Louisiana Creole, as Neumann makes explicit (1987: 20). The following question thus arises: How can we account for the difference in behavior between short and long verb stems? Rottet's analysis is that verb movement in the present tense occurs due to the presence of a null tense inflection in T, this morpheme being an affix and a trigger for V-raising (1992: 278). The long stem, in contrast, does not have any inflectional morphology, and hence no affix in T to act as a trigger for movement (ibid. 280). As a result, the long verb stem remains in situ. The case of Louisiana Creole is interesting in that it is precisely the verb stem with no overt inflection that raises to T, an analysis in sharp contrast with current assumptions that *overt* morphology triggers V-raising.

This comparison of Capeverdean, Haitian, Guinea-Bissau Creole and Louisiana Creole demonstrates the complexity of this situation, indicating that none of the theories of verb movement so far proposed have been truly able to account exhaustively for the behavior of verbs in these (and other) creoles.

6. Conclusion

The implications of the observations presented here for the theory of V-raising are twofold. First, we see how misleading the concept of "rich" verbal morphology is in predicting V-raising, in that V-raising is found in Capeverdean in spite of its having but a single inflectional marker. A more minimalist definition of the morphology required to trigger verb movement is needed.

A second obvious implication is that the explanation for V-raising may be more structural (number of heads above V^0) than morphological, as Bobaljik & Thráinsson argue. However, morphology must still play a role as a trigger, given its importance from the learnability perspective. Indeed, the Capeverdean child presumably uses V-raising because s/he is provided with the *-ba* cue, whereas the Haitian child will not, in the absence of such a cue. Nevertheless, we have seen that one of the weaknesses of this hypothesis is that it would

predict that Capeverdean would be endowed with two subject positions, one in Spec AgrP and one in TP as well as a Spec in AgrOP for shifted objects, when in fact this does not appear to be the case.¹⁸

The situation in Guinea-Bissau Creole and Louisiana Creole is, however, more ambiguous. We have yet to determine which trigger the Guinea-Bissau Creole-speaking child uses to allow the verb to raise past quantifiers. In the case of Louisiana Creole, it is difficult to reconcile overt morphology as a trigger for V-raising with the fact that in this creole, it is precisely the bare verb stems which raise, and not the full verb stems, despite their being endowed with the suffix *-e*.

If the analysis here is on the right track, then it is reasonable to assume that there are more triggers for verb movement than theories proposed so far have been able to identify.

Notes

1. Although (8b) is clearly ungrammatical (this example was provided by the editor), it is worth observing that not all utterances of a similar pattern would be starred. Indeed, consider (i), the counterpart to (8):

- (i) a. *John often goes to the movies.*
b. **/? John goes often to the movies.*

Interestingly, a number of native speakers, including the editor to this volume, would consider (ib) grammatical or would grant it at most a small question mark.

2. It is important to emphasize that only finite verbs behave this way in English. Indeed, in English, for instance, the auxiliaries *be* and *have* do raise, as illustrated in (ia) and (ib) respectively:

- (i) a. *John is not being facetious.*
b. *John has not gone home.*

3. The concepts of “weak” and “strong” have now become controversial (cf. Chomsky 1997 Fall lectures).
4. Although their formulation of verb movement is altered in more recent work, Vikner and Rohrbacher still relate verb movement to verbal morphology.
5. See also Groat (1997).
6. Jonathan Bobaljik (personal communication) has brought to my attention the fact that Afrikaans may represent such a case.
7. The generalization that all verbs, including *e*, are post-Neg holds at DS: we may assume that *e* starts out in a post-Neg position in V^0 and raises.

8. In my own idiolect and that of all the informants interviewed, it would be ungrammatical to place *e* in a post-Neg position. I was informed, however, that there seem to be dialectal varieties that use *e* in a post-Neg position, much in the same way Portuguese does.
9. *Goltarpu* is an anagram of *Portugal*. This example is taken from Manuel Veiga's 1987 novel *Oju D'Agua*. In this fictional work, the author attempts to reconcile the imaginary to the real, and refers to real places and people by using imaginary nouns (i.e. anagrams), thus allowing the reader to recognize real life characters and places.
10. As keenly observed by an anonymous reviewer, an alternative analysis would be to assume that *e* is in subject position, as it has been proposed for its Haitian counterpart *se*. For arguments pro and against this analysis, see Baptista (1999). One possible counter-argument to this hypothesis is the existence of sentences like (i), for which it would then be necessary to assume that *João* is in a topicalized (extraposed) position:

- (i) *João e malandru.*
 João e crafty
 'João is crafty.'

It is important to note that a similar construction to (i) in Haitian would yield ungrammaticality, as illustrated in (ii):

- (ii) *Bouki (*se) malad.*
 Bouki SE sick
 'Bouki is sick.' (DeGraff 1995: 237-256)

Indeed, in Haitian, adjectival predicates (as well as prepositional and many bare nominal predicates, as shown in DeGraff 1995) do not contain an overt copula.

11. The IP structure is used in (34) and (35) for ease of exposure.
12. It is worth mentioning that according to our definition, whether the verb is stative or nonstative, the verb still moves. Consider (i) where the stative verb *konxe* 'to know', occurs in a pre-adverbial position:
 - (i) *João konxe mal es país.*
 João knows bad this country
 'João knows this country badly.'
13. I thank Alain Kihm for his assistance on this issue.
14. It is important to note that (52) has a close counterpart in some dialectal varieties of Capeverdean where *txeu* 'a lot', would be preferred to *mutu* which is closer to Portuguese. In this case, both (ia) and (ib) are grammatical:
 - (i) a. *João kre Eliza txeu.* (Capeverdean Creole)
 João loves Eliza a lot
 'João loves Eliza a lot.'
 - b. *João kre txeu Eliza.*
 João loves a lot Eliza
 'João loves Eliza a lot.'

The sentence (ib) shows that the verb can raise past *txeu*, just as it does with *mutu*.

15. Interestingly, the equivalent to Guinea-Bissau *diritu* also exists in Capeverdean and is *dretu* which corresponds to *ben*. If one replaces *ben* with *dretu*, the following sentences obtain:

- (i) a. *João xina se lison dretu.* (Capeverdean Creole)
 João learned his lessso nwell
 ‘João learned his lesson well.’
 b. *João xina dretu se lison.*
 João learned well his lesson
 ‘João learned his lesson well.’

Again, in contrast with the Guinea-Bissau Creole verb, the Capeverdean verb can raise past the VP-adverbial.

16. I thank Alain Kihm for these valuable data.
 17. It is important to note that not all verbs show such an alternation in Louisiana Creole.
 18. I am currently working on a manuscript where I show evidence for 2 subject positions in Capeverdean Creole.

References

- Baptista, M. 1997. *The Morpho-Syntax of Nominal and Verbal Categories in Capeverdean Creole*. Harvard University PhD dissertation.
 Baptista, M. 1999. “On the nature of the morpheme *e* in Capeverdean Creole: To be or not to be”. In *Lenguas Criollas de base lexical Española y Portuguesa*, K. Zimmermann (ed.), 25-47. Frankfurt: Vervuert.
 Belletti, A. 1990. *Generalized Verb Movement: Aspects of Verb Syntax*. Turin: Rosenberg & Sellier.
 Bobaljik, J. 1995. *Morphosyntax: The Syntax of Verbal Inflection*. MIT PhD dissertation, MIT.
 Bobaljik, J. and Thráinsson, H. 1997. “Two heads aren’t always better than one”. *Syntax* 1: 37-71.
 Chomsky, N. 1993. “A minimalist program for linguistic theory”. In *The View from Building 20*, K. Hale and J. Kayser (eds), 1-52. Cambridge: MIT Press.
 DeGraff, M. 1995. “On certain differences between Haitian and French predicative constructions”. In *Contemporary Research in Romance Linguistics*, J. Amastae, G. Goodall, M. Montalbetti and M. Phinney (eds.), 237-256. Amsterdam: John Benjamins.
 DeGraff, M. 1996. “Verb syntax in, and beyond, creolization”, MIT ms.
 DeGraff, M. 1997. Verb syntax in, and beyond, creolization. In *The New Comparative Syntax*, L. Haegeman (ed), 64-94. London: Longman.
 Gray, D. 1985. *The Oxford Book of Late Medieval Verse and Prose*. Oxford: Oxford University Press.
 Groat, E. 1997. *A Derivational Program for Linguistic Theory*. Harvard University PhD dissertation.

- Jonas, D. and Bobaljik, J. 1993. "Specs for subjects: The role of TP in Icelandic". In *Papers on Case and Agreement I: MIT Working Papers in Linguistics Vol.18*, J.D. Bobaljik and C. Philips, (eds.), 59-98.
- Kihm, A. 1994. *Kriyol Syntax: The Portuguese-based Creole Language of Guinea-Bissau*. Amsterdam: John Benjamins.
- Koopman, H. 1984. *The Syntax of Verbs: From Verb-movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris.
- Neumann, I. 1985. *Le Créole de Breaux Bridge, Louisiane: Étude Morphosyntaxique, Textes, Vocabulaire*. Hamburg: Helmut Buske.
- Neumann, I. 1987. *Textes Anciens en Créole Louisianais*. Hamburg: Helmut Buske.
- Platzack, C. 1988. "The emergence of a word order difference in Scandinavian subordinate clauses". In *McGill Working Papers in Linguistics: Special Issue on Comparative Germanic Syntax*, 215-238.
- Pollock, J-Y 1989. "Verb-movement, UG and the structure of IP". *Linguistic Inquiry* 20: 365-424.
- Roberts, I. 1993. *Verbs and Diachronic Syntax: A Comparative History of English and French*. Dordrecht: Kluwer.
- Rohrbacher, B. 1995. "Explaining the syntactic consequences of 'rich' agreement morphology in the licensing of V-to-AgrS raising and pro". In *The Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, R. Aranovich, W. Byrne, S. Preuss and M. Senturia (eds.), 350-364.
- Rottet, K. 1992. "Functional categories and verb movement in Louisiana Creole". *Probus* 2: 261-289.
- Sportiche, D. 1988. "A theory of floating quantifiers and its corollaries for constituent structure". *Linguistic Inquiry* 19: 425-449.
- Thráinsson, H. 1996. "On the non(-universality) of functional categories". In *Minimal Ideas: Syntactic Studies in the Minimalist Framework*, A. Werner, S. D. Epstein, H. Thráinsson and C. J. Wouter Zwart (eds.), 253-281. Amsterdam: John Benjamins.
- Veiga, M. 1987. *Oju d'Agu*. Praia, Cape Verde: Instituto Caboverdiano do Livro.
- Vikner, S. 1995b. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.

Notes on Componential Diffusion in the Genesis of the Kabuverdianu Cluster

Angela Bartens¹

1. Introduction

Kabuverdianu is a Portuguese-based creole language spoken by more than a million Cape Verdeans.² Due to adverse ecological and economic conditions, only about one third of all Cape Verdeans live in the Cape Verdean archipelago, while the other two thirds constitute a culturally vital diaspora. This diaspora is mainly represented in the mother country, Portugal, and in the New England region of the United States, especially Massachusetts. In fact, Massachusetts has had a Cape Verdean community for more than a hundred years and the population of this community is now almost equal to that of the Cape Verdean archipelago itself; São Tomé e Príncipe, the Netherlands, Italy, and Senegal have been other important targets of the Cape Verdean emigration in the recent past.

The nine inhabited islands of the Cape Verdean archipelago are divided into the four Sotavento Islands, Santiago, Fogo, Maio, and Brava, and the five Barlavento Islands, Santo Antão, São Vicente, São Nicolau, Sal, and Boavista. Santiago is the most populous island with over 170,000 inhabitants, 62,000 of whom are in the capital Praia, the political center of the archipelago. Due to its natural harbor, Mindelo, on the island of São Vicente, represents a link to the outside world and continues to be regarded by most as the cultural capital of the Cape Verde Islands.

2. The Genesis of the Kabuverdianu Dialect Cluster

The southern and eastern islands of the Cape Verdean archipelago were

discovered by the Italians Cadamosto and Noli in 1456 and 1460, respectively.³ Subsequently, the islands were claimed by the Portuguese crown, but only Santiago and Fogo were settled, with Portuguese from the Algarve and Alentejo regions. These peninsular provinces had been colonized only shortly before, in the aftermath of the Reconquista, a fact which is likely to have left its imprint on the dialects spoken there in the form of some degree of koineization. Other population groups included Jews, Crypto-Jews, Marrans, and some Spanish, Italians, and Flemish, and also slaves from the African mainland opposite, who thus were most likely speakers of West Atlantic and Mande languages.⁴

In 1513, on Santiago and Fogo there were 130 whites, including 4 women, as well as 16 male and 16 female slaves. By 1582, the adult population had surpassed 15,000, but Europeans and the free colored constituted no more than 12.8% (Carreira 1982: 44, 59-61). Fanha (1987: 293) specifies that there were 13,700 slaves and “about a hundred” whites; this leaves us with over 1200 free colored. Carreira (1982: 60) distinguishes a number of groups who, at least on the island of Santiago, divided both the geographic and the social space in a way which must have contributed to the genesis of Kabuverdianu, especially outside the capital Ribeira Grande. The groups identified by Carreira are: white *vizinhos* (lit., ‘neighbors’) with a licence from the King (708 on Santiago, 300 on Fogo); *escravos de confissão* (‘slaves of faith’; 9500 on Santiago, 1500 on Fogo) who mainly lived in Ribeira Grande on Santiago; *brancos e pardos*, whites and persons of mixed descent (600); *pretos fôrros* (‘freed slaves’; 400); and *escravos em doutrinação* (‘slaves undergoing indoctrination’, 2200 on Santiago, 500 on Fogo). The *brancos e pardos* and the *pretos fôrros* were only represented in the interior parishes of Santiago; the *escravos em doutrinação* lived in the interior parishes of Santiago as well, and also in Praia (which did not become the capital until 1858) (cf. Lang 1994).

By the end of the 17th century, plantations had proved unprofitable and, since Portuguese ships had obtained permission to sail directly from Senegambia to the Americas, the profitable triangle trade had lost its importance.⁵ At this point in the archipelago’s history, there were only about 20 whites remaining on Santiago (Fanha 1987: 294). However, a little over a century later, in 1807, the population of the archipelago consisted of 1752 whites, 5139 slaves and 51,540 free colored (Carreira 1982: 61-62). The emancipation of slaves was thus a reality long before abolition, a development which presumably also left its imprint on the creolization outcomes in the

Cape Verde Islands.⁶

Meanwhile, as a rule, the Barlavento Islands were settled much later and with a greater proportion of whites than were present in the Sotavento Islands. Furthermore, the Europeans on these islands frequently immigrated from other islands in the archipelago. The eruption of the Pico de Fogo in 1500 led to an exodus of the settlers on that island, which had been settled from 1480 onward, to neighboring Brava. There must have been only a few slaves on Fogo at this early date (in 1513, Santiago and Fogo had a population of 130 whites and 32 slaves altogether), and the subsequent isolation of the island led its population to remain quite light-skinned. This very isolation forced the Bravaenses to be among the first to board whaling ships and to settle in New England in the 19th century. The isolation has continued until recent times; for example, the island's airport was completed only in 1991. As a consequence, the creole that is spoken on Brava is at least mesolectal in terms of a hypothetical Kabuverdianu continuum.

For several centuries beginning in 1497, Boavista remained in the hands of the family of the pioneering navigator Diogo Afonso; this island has a high proportion of whites even today, which is reflected in the fact that the local Kabuverdianu variety seems to be the one in which creolization proceeded to the weakest extent in the archipelago. The island of São Nicolau, discovered in 1461, was inhabited only by slaves and herds of goats, cows, and horses during the 16th century, and it was not until the 17th century that the first white settlers from Portugal and Madeira and also from the other islands arrived. Sal was tentatively settled by Europeans and slaves from São Nicolau and Boavista in the course of the 18th century; larger-scale settlement on this island dates from the early 19th century (from 1808 onwards). Like Boavista and Sal, Maio became a matter of interest only because of the salt it produced. Until the 17th century, only a relatively small population, consisting mainly of slaves and herds of domestic animals, inhabited the island. São Vicente was settled in 1794 with whites from the other islands, the Azores, Madeira, and Portugal, and also a few slaves. After massive exploitation by feudal lords beginning in 1548, exploitation by the Companhia Geral do Grão-Pará from 1757 onwards, and the effects of devastating droughts, Santo Antão had to be virtually repopulated in the 19th century. There were some wealthy Jewish families from Morocco among the white settlers at that time (as was also the case for the other Barlavento Islands) who contributed to Santo Antão's economic revival.

After unsuccessfully trying to install a plantation economy on Santiago, the Portuguese realized that the other islands were even less likely to permit more than small farms which were just big enough for subsistence farming and for supplying the ships that passed this strategically important post in the Portuguese empire. Thus, the more arid eastern islands, including the Sotavento island of Maio, were first regarded as reservoirs for meat supplies. A handful of slaves were left on the islands to keep an eye on the roaming herds of goats, cattle, and horses, the latter of which turned out to be a valuable currency used to obtain slaves on the African mainland (see above).

The dichotomy of field vs. domestic slaves seems an oversimplification of many complex situations, ones which should be separately defined in order to understand each creolization process (cf. Singler 1990, Bailey, Maynor & Cukor-Avila 1991). The interaction between masters and slaves was much closer on the Cape Verde Islands than in the classical creolization situation defined in Bickerton's Pidginization Index (1984); for example, Carreira (1982: 44, 59) stresses the role of the *escravos de casa* ('domestic slaves') in the formation of Kabuverdianu.

In comparison to the Spanish, the Portuguese constituted a very small minority in their colonies, the population of their mother country being much smaller to start with (1 million at the beginning of the Portuguese colonization as opposed to 7 million Spaniards at the beginning of the Spanish colonial expansion [Bartens 1996: 21, fn. 13]). Furthermore, there was a significant disproportion between the sexes, the bias being towards male immigrants from Portugal. This led to a mixing of races in the Portuguese colonies which was perhaps most thorough on the Cape Verde Islands, where the tendency towards the total elimination of the white and black ends of the racial continuum persists. At present, the average figures for the archipelago are 3% whites, 23% blacks, and 74% mixed. However, over 50% of the population of the main island of Santiago are black while 90% of the population of São Vicente are mixed, a fact which explains why the latter group finds it difficult to identify with the creole of Santiago. Occasionally, both Boavista and Brava are quoted as being the "white islands" of the archipelago (see below).

Given the above, it seems probable that creolization was both early and rapid on Santiago. On the other hand, as noted earlier, there was an uninterrupted supply of slaves during the initial period of settlement. Until about the 1640s, the slaves were first shipped to the Santiago entrepôt. After christianization and ladinization, most of the slaves were then shipped on to destina-

tions in the Canary Islands, on the Iberian Peninsula, and in the New World. This constant influx of new slaves guaranteed a continuous repidginization until well into the 17th century, and ultimately led to the coexistence of different varieties representing different degrees of creolization (cf. Fanha 1987). The frequent use of Cape Verdean *línguas* (interpreters) by the Portuguese in the early exploration of the West African coast hints at bilingualism or even trilingualism in creole, Portuguese, and African languages for at least part of the population, since newly imported *boçais* could not have been used in this task given their lack of competence in Portuguese, creole, or both. Additionally, there is evidence that whites used the creole in their correspondence from early on (cf. Carreira 1982, Fanha 1987: 296-297, Thiele 1991a: 28-29); this is in accordance with the language policy the Portuguese practiced throughout their sea-borne empire, and which resulted in a linguistic infrastructure that was later taken over especially by the Dutch, such as in Ceylon.

By combining the diffusion model of creolization as outlined in Holm (1986), together with the notion of the importance of the proportions of the social groups or components present during each creolization process as suggested by Hancock (1986),⁷ we obtain the componential diffusion model which I have argued for earlier (Bartens 1996). The central idea of this proposed model is that of monogenesis from an Afro-Portuguese pidgin and/or a comparable reduced variety of a European language (i.e., transplantation), both direct and indirect substrate influence, universal L2-acquisition strategies, constraints of Universal Grammar in general and, last but not least, superstratal, often dialectal, structures, all of which would have contributed to the emerging creole. Within this model, the contribution of each of these factors would vary, depending on the circumstances of the particular creolization situation, including the type of economy and society and especially the numerical and social relations between the different population groups which would determine the type of linguistic input individual speakers would receive. In the case of Kabuverdianu, different diachronic and possibly also diatopic (geographical) varieties of Portuguese seem to have been involved, too.

Although racial mixing and the ensuing social mobility were particularly widespread in the Cape Verde Islands as a whole, notable differences existed among the various islands. For example, there was labor-intensive agriculture and greater social distance at least during the initial period on Santiago as opposed to the generally closer social interaction that existed on many

Barlavento Islands (see above). The genesis of the Kabuverdianu dialect or even language cluster could therefore be considered a fairly good example of the functioning of the diffusion model.

The diffusion model might be used to explain a number of the features of the creolization process in the Cape Verde Islands. First, globally speaking, the outcomes of creolization in the Barlavento Islands were less distant from superstratal structures than those in the Sotavento. This is seen in passages of Barlavento creole which mix Portuguese and creole, a result of inability or lack of motivation on the part of speakers to distinguish between the two codes which is apparently more frequent on the Barlavento Islands. Many Cape Verdeans maintain that only Santiago, and possibly Fogo and Maio, possess a genuine creole (cf. the evaluation of the dialect of Santiago by inhabitants of other islands reported as early as 1841 quoted below). Thiele (1991a: 34) quotes a study by Scotti-Rossin (1983) which reports that Portuguese was felt to be a foreign language without a transparent genetic relationship to Kabuverdianu by speakers of the São Vicente variety.

Second, the other varieties of creole, including the Barlavento creoles, must be regarded as dilutions of the original creole of Santiago. By this we mean that the creole of Santiago was transplanted as part of the superstrate to the other islands, where a (partly) new creolization process took place so that both mono- and polygenesis could be evoked to describe the genesis of the Kabuverdianu cluster.

Third, although Costa & Duarte (1967 [1886]: 239), at the end of the 19th century, maintained that on Maio, São Vicente, Santa Luzia (an island in the Barlavento where there are now only temporary settlements), and Sal there was no autochthonous creole, we have reason to believe that ultimately, separate varieties may have arisen on all of the inhabited islands (cf. Lang 1991: 4). All of the other varieties would hence constitute *second generation varieties*, as opposed to the *first generation varieties* (in the terminology of Chaudenson 1992 on Santiago).⁸

Thus we find that there are more basilectal varieties of creole in the Sotavento Islands and more acrolectal varieties in the Barlavento Islands. Contrary to previous assumptions, research by Cardoso (1989: 17-8) indicates that the variety of São Nicolau cannot be attributed to the Barlavento Islands, but apparently occupies an intermediate position between the two island groups.

The investigators Chelmicki and Vernhagen (cited in Carreira 1982: 70) testified as early as in 1841 to the existence of separate varieties of Kabu-

verdianu, with Santiago possessing the most basilectal variety:

“Todas as ilhas têm a sua corruptela diversa; pior é o de Santiago, chamado até pelos outros insulanos crioulo cerrado” (‘All islands have their own corrupt variety of Portuguese; the one of Santiago is the worst and even the inhabitants of the other islands call it deep creole’; translation mine).

Lopes (1967 [1941]), too, reports as many separate varieties as there are inhabited islands and stresses the role of extra-linguistic factors. Rougé (1994) regards the Kabuverdianu dialects of Santiago and Fogo as antecedents to the other dialects and provides an example of his postulated plurigenesis of Upper Guinea Creole Portuguese using comparative data from Santiago Kabuverdianu and from different varieties of mainland Kriol.

The existence of a separate variety on São Vicente is confirmed by its inclusion in Veiga’s grammars (1982, 1995). The cultural importance of Mindelo must have contributed to both the presumed and the genuine autonomy of this variety.

Admittedly, the linguistic distance between the varieties does not justify speaking of separate languages in the sense of “Abstandssprachen” as defined by Kloss (1967), but neither is the term “dialect” an adequate expression, since it does not convey the fact that the Kabuverdianu varieties are products of separate creolization processes.

Veiga (1995: 29) expresses a diametrically opposed point of view by maintaining that all Kabuverdianu varieties share a common deep structure:

“E isto significa que em Cabo Verde não há nove crioulos como alguns ingenuamente afirmam, mas um único Crioulo, o qual actualiza-se em diversas variantes dialectais”

(‘And this means that there are not nine creoles on the Cape Verde Islands as some naively pretend but a single Creole which is realized as different dialectal variants’; translation mine).

We do not agree with Thiele (1991a) that Kabuverdianu is a decreolized variety, a “Spätkreol”, as she calls it. We believe the variation on the basilect/mesolect/acrolect axis dates from the beginning of creolization, while the proportions of various speaker groups may shift with time and with increasing social mobility. This is how Alleyne (1980), Bickerton (1984), Baker (1990), and Chaudenson (1992) characterize the early continuum situation, especially in the former British Caribbean. Furthermore, already at the end of the 19th century, well before the social changes that might have brought about decre-

olization in Kabuverdianu in the recent past, one of the pioneers in creole studies, Adolfo Coelho, observed the existence of a distinct *crioulo rachado/fundo/velho* ('split/deep/old creole') on Santiago (1967: 5). Most recently, Rougé (1994: 143) also indicates that the sociolects of Kabuverdianu date from creole genesis. On the other hand, the last attribute given to this lect is not uninteresting: it seems to have been speakers themselves who qualified the basilectal variety as *velho* 'old'.

3. Present Sociolinguistic Situation and Language Policy

Thiele (1991a: 28, 35) characterizes the sociolinguistic situation of the Cape Verde Islands as functional bilingualism with about 40% of the population using Portuguese as a second language; the percentage of those who speak or understand Portuguese is 70-80% (Nunes 1991). Thiele's characterization seems to be quite an optimistic evaluation of the situation, which differs from classical diglossia due to the heavy interaction between Portuguese and the creole, and inherent tendency of creoles to adapt and even relexify according to the communicative needs of their speakers. These properties significantly contribute to the vitality of the creole but also further the shift to acrolectal variants (see below). Speakers distinguish between the urban *crioulo levinho* employed in slightly more formal situations and the *crioulo fundo* used in the countryside and in familiar usage. Veiga (1995: 29-35) regards the present situation as diglossic and considers bilingualism a future goal.

Only detailed dialectological research could clear the doubts concerning the extent of the geographic variation. Meintel (1975: 237-238) indicates that even on Brava, the island with the smallest surface, there is diatopical variation which has not generally been assumed to be the case for any of the islands. Upon closer examination, however, this seems to be above all a case of diastratic (social) variation, since Meintel characterizes the language of the island capital Vila Nova Cintra as acrolectal in contrast to the rest of the island.

For the time being, there seem to be neither the financial means nor the political willpower to fully recognize the linguistic diversity which exists in the archipelago. In the initial euphoria of independence from Portugal in 1975, Kabuverdianu was declared the national (albeit not official!) language of Cape Verde and it was proposed at the Mindelo colloquium in 1979 that the variety

of Santiago be introduced into the media, education, and the socioeconomic and sociocultural spheres (Veiga 1982: 13; Thiele 1991a: 30-35).

The choice of this particular variety was highly political: Santiago and the capital Praia not only constitute the political and demographic centers of gravity, but the Sotavento varieties of the creole are also closer related to the Kriôl spoken in Guinea-Bissau. Until the *coup d'état* of 1980, which was at least in part motivated by fear of cultural and political domination by the Cape Verdeans, Guinea-Bissau was to form a state union with Cape Verde. The creole of Santiago can also be regarded as the most basilectal, and is therefore considered by many Cape Verdeans as the most African of all Kabuverdianu varieties. This has to be seen in connection with the fact that independence from Portugal brought about an appreciation of the African heritage of Cape Verdean culture.

It will probably never be possible to uncover the exact genetic relationship between the Upper Guinea Portuguese Creoles, that is, whether there was transplantation from the islands to the continent, parallel development, or a combination of both. However, it is clear that the contact was extensive since the two colonies formed an administrative unit until 1879, and many of the *lançados* who first penetrated into Guinea came from the Cape Verde Islands. It seems likely that there was mutual influence between the islands and the mainland with Cape Verde dominating, while separate creolization processes took place on the islands as well as on the mainland. A linguistic continuum of Portuguese — Barlavento varieties of Kabuverdianu — Sotavento varieties — Kriôl has also been postulated (cf. Morais-Barbosa 1967: 138, Cunha 1981: 43, Perl 1982: 65-66, Fleischmann 1984: 128-129, Pereira 1991: 20, Thiele 1991a: 37-38, Rougé 1994).

The *coup d'état* of 1980 in neighboring Guinea-Bissau, mentioned earlier, also slowed down the pro-creole language policy of the Cape Verdean government. In addition the elite of Mindelo would not accept linguistic domination by the presumedly more African and backward and, hence, stigmatized population of the Sotavento Islands, nor would they accept the phonemic orthography devised by Veiga (e.g. 1982: 13-14) upon the recommendations of the Mindelo colloquium. Later, all varieties were to receive a separate orthography which, however, proved too expensive. For several years, little concrete progress visible to the outside world was made although the Comissão Nacional para a Língua Cabo-Verdiana ('National Commission for the Cape Verdean Language') and the Forum de Alfabetização Bilingue

(‘Forum for Bilingual Alphabetization’) were created in 1989, and the Comissão Nacional para a Padronização do Alfabeto (‘National Commission for the Standardization of the Alphabet’) was created in 1993 (at the time of the Colloquium of Mindelo, the existence of a phonematic alphabet by Ana de Paula de Brito [1967[1887]] was ignored [Veiga 1995: 26-27]). Recently, pro-creole language policy seems to have picked up and a new orthography (ALUPEC, Alfabeto Unificado para a Escrita do Crioulo, ‘Unified Alphabet for Writing Creole’), a compromise between a phonemic orthography and traditional etymologizing spelling systems, is awaiting approval by the Government. Approval of the new orthography would be an important measure in that it would pave the way toward further institutionalization and ultimately, toward the maintenance of creole together with Portuguese.⁹ On the other hand, the extensive linguistic accommodation during speaker-interaction observed by Fanha (1987) is typical of creole speech communities and does not in itself testify to decreolization or language shift.

4. Dialectological Studies on Kabuverdianu

Obviously the emergence of a uniform creole would boost nation-building, a task which has not always been easy in this heterogenous island community. This might take place through official promotion of one variety or through natural selection; the latter is a somewhat less likely prospect, although it was at one time advanced by the government (cf. Fleischmann 1984: 132, Holzer 1991).¹⁰ We must note the observation of Batalha (1985) that the language death which mobilizes linguists from all over the world may actually mean economic opportunities and social promotion for the community of speakers who are at least stigmatized, or at most excluded from mainstream society because of their divergent speech. In many cases, bilingualism is the best solution (in contrast to official monolingualism and diglossia) and it is the one which Veiga advocates for Cape Verde (1995: 29-33; cf. also Fishman 1991: 60-61, 64-65).

However, a precondition for both status and corpus planning is the study and adequate documentation of the linguistic variation in the Cape Verde Islands, an undertaking which would simultaneously result in recording part of their cultural heritage. It is also important to systematically compare the different Kabuverdianu varieties to Portuguese dialects since at least part of

the divergences might be accounted for by different dialectal superstratal input. All future research must be by or in cooperation with the Cape Verdean people themselves. Indeed, the local media have appealed to the population to collect folklore, and contemporary Cape Verdean linguists, especially Manuel Veiga and Eduardo Augusto Cardoso, are doing dialectological research. Ever since the pioneering work by Coelho, Schuchardt, and Vasconcellos, the variety of Santiago has frequently been taken as the point of reference, a procedure which thus distorts the analyses of the other varieties. In fact, the Santiago dialect often constitutes the only object of description, a tradition continued, e.g., by the German creolists Lang and Thiele. However, descriptions of the other varieties are not lacking and also date from the beginning of Cape Verdean studies: Costa & Duarte (1967 [1886]) give the *Parable of the Lost Son* in the dialects considered autonomous by them, i.e., in the varieties of Santiago, Fogo, Brava, Santo Antão, São Nicolau, and Boavista; Silva (1984 [1957]) concentrates on the variety of São Nicolau; the data presented in Almada (1958) is primarily from São Vicente but information on the variety of São Nicolau is also included; Meintel (1975) provides the most comprehensive study thus far of the variety of Brava. Veiga (1982) is a comparative grammar of the dialects of Santiago, Fogo, São Vicente, and Santo Antão, and Veiga (1995) compares the creoles of Santiago and São Vicente with Standard Portuguese;¹¹ Cardoso (1989) offers a detailed description of the variety of São Nicolau, which is probably the first generativist study in Cape Verdean studies. The texts reunited in Parsons (1923) could probably be used in the reconstruction of diatopic variation, although the possibility of dialectal leveling in the New England community would have to be taken into account, and informants from Fogo may be over-represented.¹² Nevertheless, the transcripts seem to confirm the division into Sotavento and Barlavento, with the variety of Boavista being the most acrolectal one, and São Nicolau occupying an intermediate position.

5. Comparison of Structures Within the Kabuverdianu Cluster

In this section, we present data from the different levels of the language system which support the thesis of separate creolization processes within the Kabuverdianu cluster. The choice of features in part reflects those best-suited to demonstrate differences between the dialects. For a more comprehensive

picture, see Carvalho (1962), Veiga (1982, 1995), Cardoso (1989), Thiele (1991a, 1991b), Lang (1991, 1993), and Bartens (1995).¹³

5.1. Phonology/phonetics

On the phonological level, the distribution of oral and nasal vowels, labials, sibilants, and liquids are the most salient features of the Cape Verdean creoles lending support to a dialectal division into Sotavento and Barlavento (see Tables 1 and 2).

Table 1. Phonological criteria for the dialectal divisions Sotavento and Barlavento

<i>Sotavento</i>	<i>Barlavento</i>
1. maintenance of oral vowels	reduction of oral vowels
2. denasalization of vowels	nasality of final vowels is distinctive feature
3. [v] only in recent loan words	opposition /b~v/ as in modern Standard Portuguese
4. [s] predominates	[z] in intervocalic, [S _ɰ] in final position
5. Fogo, São Nicolau: instability of etymological opposition /l ~ r/	

Table 2. Examples of the distribution of phonological variants from Veiga (1982)

Santiago	São Vicente	Portuguese	English
<i>kumesá</i>	<i>kunsâ/kmesâ</i>	<i>começar</i>	‘to begin’
<i>makaku</i>	<i>makôk</i>	<i>macaco</i>	‘monkey’
<i>dipendenti</i>	<i>dpendent</i>	<i>dependente</i>	‘dependent’
<i>bon</i>	<i>bõ</i>	<i>bom</i>	‘good’
<i>baka</i>	<i>vaka</i> (?)	<i>vaca</i>	‘cow’
<i>vérbu</i>		<i>verbo</i>	‘verb’
<i>kasa</i>	<i>kaza</i>	<i>casa</i>	‘house’
<i>ántis</i>	<i>antš</i>	<i>antes</i>	‘before’
<i>dipos</i>	<i>dpož / dpoš</i>	<i>depois</i>	‘after’
Fogo: <i>argen</i> <i>markriadu</i>	(<i>žent</i>) <i>malkriôd</i> (<i>gente</i>)	<i>alguém</i> <i>malcriado</i> (people)	‘somebody’ ‘bad-mannered’
Santiago: <i>algen</i> <i>malkriadu</i>			

More specifically, the different realizations result from: (1) different degrees of restructuring, for instance, the instability of liquids which we find especially on Fogo and on São Nicolau, is typical of basilectal Atlantic creoles and considered an L1-acquisition/substrate feature; and (2) the fact that genesis of the Kabuverdianu varieties took place at different times. During the formative period of the Sotavento (but not Barlavento) creole, the modern (Southern) Portuguese opposition /b ~ v/ was still realized as [b ~ β], a difference which was harder to perceive for the slaves who were acquiring Portuguese as a second language. This difficulty therefore resulted in the neutralization of the opposition as /b/; furthermore, it should be noted that, across languages in general, [β] is not a common sound.¹⁴ Usually, the two factors mentioned, i.e., different degrees of restructuring and genesis at different points in time, complement each other. For instance, the conservation of oral vowels and, as a consequence, the CV syllable structure, date from the Portuguese anterior to the 18th century, and are also typical of basilectal creoles. The Sotavento creoles participated in the reductions /e/ > [i], /o/ > [u] which took place in Portugal during the first half of the 18th century and which are commonplace in Portuguese-based creoles and Brazilian Portuguese. In contrast, the Barlavento varieties reflect the later stages of the vowel reduction processes in European Portuguese (cf. Almada 1958: 51, Bartens 1995, 1996). Nasality of final vowels in polysyllabic words is a relevant feature in the Barlavento dialects as it is in Portuguese, while overall denasalization of vowels has taken place in the more strongly restructured Sotavento varieties. The restructuring of the sibilant system of European Portuguese started in the 17th century and is reflected in the Barlavento creoles, while the collapse of the sibilants into /s/ in the Sotavento creoles is also due to internal L2-acquisitional and substrate-driven restructuring.

5.2. *Morphosyntax*

On the morphosyntactic level, variation among the Kabuverdianu dialects is considerable; this is seen, for example, in the marking of the NP for number and gender and in the verbal system. Once again, more superstratal (albeit sometimes fossilized) structures are conserved in the Barlavento varieties while restructuring has been much more profound in the Sotavento creoles.

In the creole of Santiago, plurality is marked only once on the entire noun phrase with a numeral or the superstratal suffix /-(i)s/ which is, in most cases,

attached to Spec-NP when plurality cannot be inferred from the context (as also found in Popular Brazilian Portuguese), as in, e.g., *ũa fiју* ‘my child’ vs. *ũa fiјus* ‘my children’. The same goes for enumerations:

- (1) *Grandi-s y pikinoti-ø, tudu-ø kòre pa da bason.*
 adult-PL and child all run for give look
 ‘The adults and the children, everyone ran to have a look.’ (Veiga 1986: 34)
 (cf. Port. *Os grandes e os pequenotes, todos correram para dar uma olhada.*)

Considering this structure as basic, Veiga characterizes the “double” marking *ũa fiјus* on Fogo as a case of hypercorrection (1982: 100-101).¹⁵ However, it is quite plausible to assume that the superstratal marking on both elements of the NP was never lost during local creolization in the first place, even though the dialect of Fogo otherwise belongs to the more basilectal Kabuverdianu varieties. This is plausible since the relative basi- or acrolectality¹⁶ never affects a creole language as a whole but only affects specific (lexical) structures.¹⁷ In the Barlavento dialects, e.g., the variety of São Vicente, plural *-s* is more frequent than in the Sotavento dialects (Veiga 1982: 140).

As a rule, there is no morphological genus marking on noun phrases in the Sotavento varieties. Gender is expressed lexically via postposition of Portuguese *macho* (‘male’) or *fêmea* ‘female’, as seen in the following translations of ‘sister; brother’:

- (2) a. *armun maçu; armun fémia* (Santiago; Veiga 1982: 71)
 b. *irmon; irman* (São Vicente; Veiga 1982: 140)
 (cf. Port. *irmão; irmã*)¹⁸

or by means of lexicalized pairs such as *pai/mai* ‘father’/‘mother’, *mininu/minina* ‘boy’/‘girl’ (Santiago; Veiga 1982: 101). On the other hand, morphological genus marking as in Portuguese is quite frequent in the Barlavento creoles. The following are examples from the variety of Santo Antão:

- (3) *mîer bnita*
 ‘the/a beautiful woman’ (Veiga 1982: 71)
 (cf. Port. *a/uma mulher bonita*)
- (4) *mnina malkriada*
 ‘the/an ill-mannered girl’ (Veiga 1982: 71)
 (cf. Port. *a/uma menina malcriada*)

Veiga attributes this morphological marking, especially in the variety of São Vicente, to the influence of Portuguese, i.e., to decreolization; we can only reiterate our adherence to the point of view expressed above (cf. Veiga 1982: 71).

Within the verbal system, there is on first glance a striking difference between the Sotavento and Barlavento varieties as far as basic form is concerned. In the Sotavento creoles, the basic verb forms are paroxytonic, while in the Barlavento varieties they are oxytonic. For this reason, Meintel (1975: 216) derives the Sotavento forms from the 3rd person singular of the present tense of Portuguese, and derives the Barlavento forms from the infinitive.

However, it may be possible to derive all of the Kabuverdianu verb stems from the Portuguese infinitive (except certain high frequency verbs which, as in all Portuguese-based creoles, are usually conserved in the 3rd person singular of the present tense). This derivation is plausible if we consider the possibility of there being tonal oppositions in Sotavento creole, as claimed by Pires & Hutchinson (1983), and consider that stress in European languages is frequently reinterpreted as a high tone when speakers of a tone language adapt base-language lexical items into a creole. In this process, stress has secondary importance and falls on a different syllable in the emerging creole (cf. Römer 1991; Devonish 1989).

However, in the verbal marking system, too, the Barlavento creoles conserve many more superstratal structures than do the Sotavento creoles. This is seen in the fossilized subjunctive forms of Portuguese auxiliary verbs. For example, the creole of São Vicente conserves the third person singular conjugated forms *ê, era, foi, fos, for* of the present, imperfect, perfect indicative, imperfect and future subjunctive respectively of the Portuguese copular verb *ser*, as well as the infinitive form, while the creole of Santiago conserves only *ê, era* and *ser* and has created *serba* by suffixing the past marker *-ba*.

Another superstratal conservation is passive constructions, as in, for example:

- (5) *Es trabòj foj / ten sid fejt pa min.* (São Nicolau)
 this job was have been made for me
 'This job was done by me.' (Cardoso 1989: 67)
 (cf. Port. *Esse trabalho foi / tem sido feito por mim.*)

The Sotavento creoles, on the other hand, have undertaken considerable innovations, as in the following mediopassives:

- (6) *Ka ta papiadu na mésa.* (Santiago)
 NEG HAB talk LOC table
 ‘One does not speak at the table.’ (Veiga 1982: 157)
 (cf. Port. *Não se fala à mesa.*)
- (7) *Na nos kasa kume-da miçu na anu pasadu.* (Santiago)
 LOC our house eat-SUF corn LOC year past
 ‘At our home, people ate corn last year.’
 (cf. Port. *Na nossa casa, comeu-se milho no ano passado.*)

Lang claims the suffix *-da* is a fusion of the other postverbal Kabuverdianu affixes *-du* and *-ba* (1993:158-160). The form ending in *-du* is of course the Portuguese past participle. The verbal action marked with the suffix *-da* is both anterior to an action marked with the suffix *-du* and is marked as being in irrealis mode. The Barlavento creoles do not share this innovation: for example, sentence (7) above translates into the variety of São Vicente as:

- (8) *Na nōs kaza žent tiña kmid miç na an’ pasód.*
 LOC our house people have eaten corn LOC year past
 ‘At our home, people ate corn last year.’ (Veiga 1982: 157)
 (cf. Port. *Na nossa casa, a gente tinha comido milho no ano passado.*)

Contrary to Thiele (1991a), we believe some of the innovative mediopassive uses are latent in Portuguese as well, while their range has increased as a result of the complete loss of the gerund in the Sotavento varieties, for example:

- (9) *Nu sta moradu na un kasa ki nu gosta,* (Santiago)
 we CONT living LOC a house REL we like
mas k’e ka nos ki fase...
 but REL-COP NEG we REL make
 ‘We live in a house which we like but which we haven’t built ourselves...’ (Silva 1988: 54)
 (cf. Port. *Estamos morando numa casa de que gostamos mas não fomos nós que a construímos...*)

Cardoso (1989) identifies four moods in the creole of São Nicolau: declarative, hypothetical, injunctive, and eventual. In contrast, Veiga (1982), agreeing with Santos (1979), lists only three moods: annunciative, eventual, and injunctive. He does not consider the cases in which Portuguese subjunctive

forms were conserved as representative (cf. Veiga 1982: 159). However, later, in his 1995 grammar, which orients itself much more closely to Portuguese grammaticography, Veiga recognizes the existence of four “modalidades modais” (‘modal moods’): indicative, conditional, imperative, and conjunctive (1995: 201-204).

In other cases, there is little controversy about the correspondence of functions, but the allomorphy is at least superficially confusing. We cite the example of the markers used to express habitual aspect/irrealis mood, continuous aspect, and anterior forms; see Table 3. According to Silva (1984 [1957]: 139), *sa* and *ta* are both reflexes of Portuguese *está* (note, however, Afro-Portuguese *sa* which is a conflation of *ser* and *estar*; e.g., Bartens [1996: 101-102]). Silva derives the particle *ti* via phonetic dissimilation from *ta* (p.c.; by analogy, this would apply to *te* as well).¹⁹

Table 3. The expression of habitual aspect (HAB)/ irrealis mood (IRR) and continuous aspect (CONT) (including anterior forms [ANT]) in four Kabuverdianu-varieties (adapted from Veiga 1982).

	Santiago	Fogo	São Vicente	Santo Antão
HAB/IRR	ta V	ta V	ta V	ta/te V
ANT/HAB/IRR	ta V-ba	ta V-ba	tá/tava V	tava V
CONT	sa ta V	sta V	ti ta V	ti te/ti ta V
ANT	sa ta V-ba	stá ta V	tá ta/	tava te/
CONT			tava ta V	tava ta V

Sentences (10), (11a, b), and (12a, b) are examples of the use of the habitual aspect, the irrealis mood, and the continuous aspect, respectively:

- (10) *N ta kume.* (Santiago)
I HAB eat
‘I eat (habitually).’ (Veiga 1982: 123, fn. 1)
(cf. Port. *Como* [*habitualmente*]).
- (11) a. *N ta kontinua konsiderasons sobri dibiña.* (Santiago)
I IRR continue considerations on riddle
‘I will continue with [my] observations on riddles.’ (Silva 1992: 26)
(cf. Port. *Vou continuar com as considerações sobre adivinhações.*)

- b. *No ta ba žunt.* (São Vicente)
 we IRR go together
 ‘We will go together.’ (Veiga 1982: 85)
 (cf. Port. *Vamos ir juntos.*)
- (12) a. *N sa ta purpara un raportaži sobri Natal.* (Santiago)
 I CONT prepare a article on Christmas
 ‘I am preparing an article about Christmas.’ (Silva 1988: 67)
 (cf. Port. *Estou a preparar uma reportagem sobre Natal.*)
- b. *Ondê k’bo ti ta ba?* (São Vicente)
 where REL-you CONT go
 ‘Where are you going?’ (Veiga 1982: 85)
 (cf. Port. *Onde é que vais?*, Braz. Port. *Onde é que você vai?*)

Sometimes several more or less synonymous constructions exist in a dialect, as in the expression of the habitual past in the creole of São Nicolau:

- (13) *N tà (ta) bà*
N tà kstuma bà(j)
N tíña kstum d’bà
N kstuma bàba
N kstuma bà
Mi èra kstumòd bà
N kstumàba bà.

‘I habitually went; I was in the habit of going’ (Cardoso 1989: 62)
 (cf. Port. *ia* (*habitualmente*); *costumava ir*; *tinha o costume de ir.*)

In the creole of Brava, Port. *costumar* ‘to be in the habit of; to be accustomed’ is also employed to form the habitual aspect (/tɒ kustúmɒ/, anterior /tɒ kustúmóbb/ [Meintel 1975: 218]). Doubtlessly, Portuguese periphrastic constructions were conserved in all Kabuverdianu varieties although they have received very little attention in the literature, cf. Soares (1947: 124) who states that “A linguagem perifrástica [...] forma-se como em português” (‘Periphrases [...] are formed as in Portuguese.’).

It seems that *-ba* is a mobile particle as in Palenquero (Schwegler 1992). The anterior of the habitual aspect and the irrealis mood is expressed by *ta* V-*ba* in the creole of Santiago and by *tá/tava* V in the creole of São Vicente (cf., Veiga 1982: 156, 158); however, the Barlavento forms are above all continuities of Portuguese *estava* (1st and, more important, 3rd person singular

imperfect indicative of the copular verb *estar*). We have reason to believe that the recruiting of the Portuguese imperfect ending *-va* as the anterior marker was influenced by the existence of a completive marker *-ban* in the Mande and West Atlantic languages (cf. Rougé 1994: 147). While the Sotavento varieties have recruited just the imperfect ending of the Portuguese *a-* conjugation-*ba* as their anterior marker, the Barlavento varieties also employ *-ia* (Portuguese *e-* and *i-* conjugations).

Apart from the rough division into Sotavento and Barlavento observable in most of the examples given so far, there are structures which only exist in one variety, e.g., the formation of the prohibitive with [ne pa] < Port. *não é para* ‘it is not for’ in the creole of Santo Antão as in [ne pa dze ɲa mãĩ] ‘Don’t tell my mother’ (Almada 1958: 213) (cf. Port. *Não digas a minha mãe*), while all the other varieties make use of the negation particle *ka*, as in [ka bo dze ɲa mãĩ] ‘Don’t tell my mother’ (cf. Port. *Não digas a minha mãe*).²⁰

The Barlavento creoles reflect the use of the copula verbs in Portuguese more closely than do the Sotavento creoles. As in 16th century Portuguese and Spanish, the distribution of the copula verbs *ser* and *estar* in the Sotavento creoles sometimes differs from modern standard Portuguese usage. However, Marlyse Baptista (1995: 14 and p.c.), a native speaker of the Brava dialect, does not judge as acceptable Meintel’s example [ɛl e dwénti] ‘He is ill’ (1975: 224) and maintains that the sentence thus glossed has to be *El sta duenti*, unless the intended reading is that of a permanent state. This parallels the distribution of *ser* and *estar* in modern Portuguese:

- (14) a. (Ele) *está doente*.
 ‘He is (currently) ill.’
 b. (Ele) *é doente*.
 ‘He is a sick person.’

At least at the present developmental stage of Kabuverdianu, zero copulas are practically nonexistent in any variety (cf. Mello, n.d.).

5.3. *Lexicon*

Apart from the preliminary dictionary by Pires & Hutchinson (1983, 1994), there is no comprehensive study of the Cape Verdean lexicon.²¹ Soares (1947: 131) indicates that the use of Anglicisms was widespread among the lower classes of Mindelo at the time of writing. These loanwords can be traced to the

role of the British in the opening of the international port of Mindelo in the middle of the 19th century. Ferraz & Valkhoff (1975: 35) take the higher proportion of Africanisms in the creole of Santiago as proof of its archaic nature.

6. Concluding remarks

As I have tried to demonstrate in the preceding sections, Kabuverdianu is far from being a homogeneous language. What we are dealing with is a cluster resulting from (partly) separate creolization processes described by the componential diffusion model. The overall scenario is further complicated by the coexistence of diatopic and diastratic variation (see above). The extent of variation can only be defined after thorough dialectological research, which would be a precondition for both status and corpus planning; it would also serve as a comparison base for the study of the emigrant communities, a field of study which promises interesting linguistic insights.

However, research would not have to begin precisely from the ground up, as the existing literature contains more data on variation than one might expect. Consequently, a first task would be systematizing this existing material and then comparing it to newly collected data; afterwards, both ought be compared with data from Portuguese dialects (see above). One of the main obstacles to such an undertaking is the heterogeneity of the existing material as far as chronology and terminology are concerned. A particular difficulty will be balancing the need for uniform terminology (and the categories behind it) against the demand for descriptions that are unbiased in terms of what is known about other Kabuverdianu varieties or creole languages.²²

Since part of its vitality results from its interaction with Portuguese, Kabuverdianu stands a good chance of surviving this external pressure. In what form it will survive, and whether it will be as a heterogeneous cluster or a uniform creole, remains to be seen and will depend last but not least on the contribution made by linguists and intellectuals.

Notes

1. I would like to thank John McWhorter and an anonymous reviewer for their pertinent comments on an earlier version of this paper. I would also like to thank Maria Carmen de Castro Duarte de Frias e Gouveia for her comments on the diachronic development of

labials in Portuguese. I am indebted to George Lang, John Ladhams, Heliana R. de Mello and Francisco de Assis Mira Espada for elucidating some of the terms used in Section 3 and/or helping me with the translation into Portuguese of some of the examples given in Section 5. Last but not least, this paper profited from the comments made by the audience after its initial presentation at the SPCL meeting in San Diego in Jan. 1996. The responsibility for all errors and shortcomings is, of course, entirely mine.

2. Speakers themselves mostly refer to their mother tongue as *crioulo* / *kriolu*. Veiga (1982: 20) recommends the use of the glossonym *Ka(b)uverdianu*. The increased need for a term which can be used in international relations and which lends expression to the national feeling has led to similar proposals in many creole speech communities (cf. Bartens, in press a).
3. Noli was most probably accompanied by the Portuguese Diogo Gomes who may even have undertaken some missions on his own; the western islands of the archipelago were discovered by the Portuguese Diogo Afonso between 1460 and 1462 (cf. Peres 1983 and Brásio 1962; the latter is desperate to prove it was a Portuguese who discovered the islands). We are referring to “discovery” strictly within the lens of Eurocentric historiography since the archipelago already appears on maps drawn by Arabian geographers at the beginning of the 15th century. The Wolof, Serere, and Lebu from Senegal came to the islands on occasion to fetch salt, and before 1288, which is the year of the death of the Arab geographer Ibn Said, there seems to have been a permanent settlement in the South of Sal (cf. Ferreira 1985: 62).
4. When the Portuguese entered the rivers of the Senegambia, all of the actual populational groups but the Fulbe had already migrated into the area (PAIGC 1974).
5. In the case of Brazil, the first permission for direct sailing routes was granted in 1644.
6. Because of high death rates due to droughts and famines, slavery was abolished already in 1780 on Santo Antão. It was abolished in 1875 on São Vicente, and in 1878 in the entire archipelago.
7. Thiele (1991b: 187) expresses a similar point of view but only refers to the linguistic input: “Der Anteil basissprachlicher, substratsprachlicher und universeller Einflüsse auf das grammatische System der Kreolsprachen muß für den Einzelfall jeweils genau überprüft werden” (‘The proportions of base language, substratal and universal influences on the grammatical system of a creole have to be carefully scrutinized in each separate case’; translation mine).
8. Of course, only the results of fieldwork will prove if each inhabited island really has its own variety. “First and second generation creole” are terms used for example by Chaudenson (1992). According to him, first generation creoles are varieties which have arisen during the initial phase of the “société d’habitation” when a relatively small number of slaves lived and worked side by side with the whites and therefore had a fairly good chance of acquiring the prestige language. These first generation creoles have usually remained fairly acrolectal but have served as input to the genesis of more basilectal second generation creoles which have typically arisen after exportation to another island/colony and, above all, after the switch to the “société de plantation” where a numerically very strong majority of slaves worked on huge plantations and had hardly any access to the standard language.

9. Diglossia is also reflected in the literary production of the Cape Verde Islands where the dichotomy of using Portuguese for intellectual and artistic topics and Kabuverdianu for popular ones dates from the 19th century (Ferreira, 1985: 257). It is quite understandable that writers should want to reach a larger audience. However, there are some contemporary authors writing in Kabuverdianu. T.V. da Silva, who has also collected folklore and who is the only representative regularly employing Kabuverdianu in Parliament writes in the variety of Santiago; the linguist Manuel Veiga also writes in the variety of Santiago; Sergio Frusoni and Corsino Fortes, who sometimes combines Portuguese and Kabuverdianu in his poems (an example of the vitality of the interaction of the two systems!), both write in the variety of São Vicente; Luis Romano writes in the Santo Antão variety. Carlos (Káka) Barbosa has published a collection of poems contrasting the old etymologic with the new phonemic spelling.
10. There are daily radio (30 minutes) and occasional TV broadcasts in creole (Lang, personal communication); the introduction of creole into the classroom is being reconsidered (Duarte 1994, Silva 1994).
11. Cf. part of Veiga's argumentation (1982: 21) in favor of choosing the variety of Santiago as the basis for creating the standard language:

“Politicamente, impõe-se tomar a variante mais representativa do ponto de vista sociocultural para facilitar a unidade nacional, alargar a intercompreensão, favorecer a produção e consolidar a reconstrução nacional. Sendo a variante de Santiago a que representa maior peso sociocultural era portanto natural que o Colóquio a recomendasse.

Entretanto, a recomendação do Colóquio não visa prestigiar uma variante em detrimento de outras.”

('It is politically imperative to choose the variety which is most representative from the sociocultural point of view in order to facilitate national unity, increase mutual comprehension, favor production, and consolidate the national reconstruction. Since the variety of Santiago is the most important one in sociocultural terms, it was natural for the Colloquium to recommend it.

However, the recommendation of the Colloquium does not intend to give prestige to one variety over another.'; translation mine.)

In addition to these arguments, which are above all political, Veiga presents sociolinguistic and linguistic arguments in favor of the variety of Santiago.

12. Veiga (1982) is also innovative in the sense that it is written almost entirely in Kabuverdianu (variety of Santiago); due to false pagination and other technical problems, Veiga (1995) was withdrawn from the market and is currently unavailable.
13. Hancock (1971: 516) indicates that decreolization is taking place in the communities in New Bedford, Massachusetts, and in California. Even when Portuguese is not the language of the receiving country, it is a powerful model for the diaspora population, transmitted, for example, through bilingual programs.
14. The collapse of the labial phonemes into [b] has a parallel in the Northern Portuguese dialects; this may have constituted a minor factor contributing to its maintenance in the Sotavento.

In his *Orthographia da Lingoa Portuguesa* of 1576, Duarte Nunes de Leão first drew attention to this difference in articulation between Galicia and Northern Portugal on the one hand, and the rest of the country on the other (Bartens, in press b).

Almada reports free or lexicalized variation from the creoles of São Nicolau, Sal, and Boavista with decreolization towards the realization [v] due to pressure from Portuguese (1958: 100-102).

15. This is precisely the kind of descriptive bias to be avoided, from which Veiga's study does not seem to be entirely free after all.
16. Limitations of space prevent me from exploring the validity of these terms; after decades of intense activity in the study of creoles, the basic concepts and terminology have not yet been defined in a satisfactory manner (cf. Winford 1997).
17. This represents evidence from creole languages for the Independence Principle as well as for the Lexical Parametrization Hypothesis as presented, e.g., in Manzini and Wexler (1987; cf. Bickerton 1988).
18. Apparently, only the marking of female nouns is obligatory (cf. Veiga, 1982: 70). Romance-based creoles have preferentially taken over nouns from the superstrate languages in their masculine form, thus considering it the unmarked member of the genus opposition.
19. Jürgen Lang cautions (personal communication) that the table adapted from the data presented in Veiga (1982) needs to be revised. Marlyse Baptista agrees (personal communication) with the analysis of Suzuki (1994) according to which *ta* has a third function, namely, marking the subjunctive mood.
20. Almada (p.c.) does not specify from which dialect the second construction is taken. The negation particle *ka* is derived from Portuguese *nunca* with possible African substrate influence (but cf. Schuchardt 1888: 217).
21. For instance, Silva (1984 [1957]) contains a section on the lexicon. At present, a dictionary of the variety of Santiago is being compiled under the direction of Jürgen Lang at the University of Erlangen-Nürnberg.
22. Immediately after the presentation of this paper in San Diego in January 1996, a handful of Kabuverdianu specialists (or specialists of other Portuguese-based creoles) met in order to make plans for dialectological studies of the archipelago. At first, the compilation of a traditional dialect atlas was proposed, last but not least because such atlases have been executed on French Creole-speaking islands. However, as time went by and particularly as an outcome of the 2nd International Colloquium on Spanish- and Portuguese-based creoles held in Berlin in October 1996, it was decided monographs on the separate dialects would better serve both scientific goals and the community, and would also constitute a more realistic project. The following linguists have agreed to participate or give their support to the project: Marlyse Baptista (University of Georgia), J. Clancy Clements (Indiana University, Bloomington), Jürgen Lang (University of Erlangen-Nürnberg), Manuel Veiga (INAC, Cape Verde), and myself. Although the project is still in its initial stages, we hope that it will contribute in a significant manner to the development of Cape Verdean studies, and even to that of the country.

References

- Alleyne, M. C. 1980. *Comparative Afro-American*. Ann Arbor: Karoma.
- Almada, M. D. de O. 1958. *Cabo Verde: Contribuição Para o Estudo do Dialecto Falado no Seu Arquipélago*. Coimbra: Faculdade de Letras.
- Bailey, G., Maynor, N. and Cukor-Avila, P. (eds) 1991. *The Emergence of Black English: Text and Commentary*. Amsterdam: John Benjamins.
- Baker, P. 1990. Off Target? [Column]. *Journal of Pidgin and Creole Languages* 5: 107-119.
- Baptista, M. 1995. "On the Nature of Pro-drop in Capeverdean Creole". *Harvard Working Papers in Linguistics* 5: 3-17.
- Bartens, A. 1995. *Die Iberoromanisch-basierten Kreolsprachen: Ansätze der Linguistischen Beschreibung*. Frankfurt: Peter Lang.
- Bartens, A. 1996. *Der Kreolische Raum: Geschichte und Gegenwart*. Helsinki: Die Finnische Akademie der Wissenschaften.
- Bartens, A. in press a. "Glosónimos, etnónimos e identidade: el glosónimo *lengua* en palenquero y otras lenguas acriollizadas". In *Palenque, Cartagena y Afro-Caribe: Historia y Lingüística*, Y. Moñino, A. Múnera and A. Schwegler (eds). Cartagena: Universidad de Cartagena and Colciencia.
- Bartens, A. in press b. "Die diasystematischen Markierungen in den portugiesischen Grammatiken und Traktaten des 16. Jahrhunderts". In *Geschichte des Sprachbewusstseins: Sprachdiskussion und Sprachbeschreibung in den romanischen Ländern*, G. Haßler and H.-J. Niederehe (eds). Münster: Nodus.
- Batalha, G. N. 1985. "Situação e perspectivas do português e dos crioulos de origem portuguesa na Ásia Oriental (Macau, Hong Kong, Singapura, Indonésia)". In *Actas do Congresso Sobre a Situação Actual da Língua Portuguesa no Mundo*, Vol. I (no ed.), 287-303. Lisboa: Instituto de Cultura e Língua Portuguesa.
- Bickerton, D. 1984. "The language bioprogram hypothesis". *The Behavioral and Brain Sciences* 7: 173-188.
- Bickerton, D. 1988. Creole languages and the bioprogram. In *Linguistics: The Cambridge Survey. Vol. II: Linguistic Theory: Extensions and Implications*, F.J. Niemeier (ed.), 268-284. Cambridge: Cambridge University Press.
- Brásio, P. A. D. 1962. *Descobrimento, Povoamento, Evangelização do Arquipélago de Cabo Verde*. Lisboa: Centro de Estudos Históricos Ultramarinos.
- Brito, A. de P. de. 1967[1887]. "Dialectos Crioulos-Portugueses. Apontamentos para a Gramática do Crioulo que se Fala na Ilha de Santiago de Cabo Verde. Revistas por F. Adolfo Coelho". In *Morais-Barbosa* (ed.), 329-404.
- Cardoso, E. A. 1989. *O Crioulo da Ilha de S. Nicolau de Cabo Verde*. Lisboa: Instituto de Cultura e Língua Portuguesa/Praia: Instituto Cabo-Verdiano do Livro.
- Carreira, A. 1982. *O Crioulo de Cabo Verde: Surto e Expansão*. Lisboa: (author's edition).
- Carvalho, J. G. H. de. 1962. "Sincronia e diacronia nos sistemas vocálicos do crioulo caboverdiano". In *Miscelânea Homenaje a André Martinet: Estructuralismo e História*. Vol. III (no ed.), 43-67. Lanuna (Canárias): Universidad de la Laguna.
- Chaudenson, R. 1992. *Des Îles, des Hommes, des Langues*. Paris: L'Harmattan.
- Coelho, F. A. 1967[1880, 1882, 1886]. "Os Dialectos Românicos ou Neo-Latinos na

- África, Ásia e América. Notas Complementares. Novas Notas Complementares". In Morais-Barbosa (ed.), 1-234.
- Costa, J. V. B. da and Duarte, C. J. 1967 [1886]. "O Crioulo de Cabo Verde. Breves Estudos sobre o Crioulo das Ilhas de Cabo Verde. Oferecidos ao Dr. Hugo Schuchardt". In Morais-Barbosa (ed.), 235-328.
- Cunha, C. 1981. *Língua, Nação, Alienação*. Rio de Janeiro: Nova Fronteira.
- Devonish, H. 1989. *Talking in Tones. A Study of Tone in Afro-European Creole Languages*. London: Karia Press/Barbados: Caribbean Academic Publications.
- Duarte, D. A. 1994. "Crioulo caboverdiano e diglossia". *Papia* 3: 26-33.
- Fanha, D. 1987. "Aspectos do contacto entre o português e o crioulo de Cabo Verde". In *Actas do Congresso sobre a Situação Actual da Língua Portuguesa no Mundo*, Vol. II (no ed.), 292-310. Lisboa: Instituto de Cultura e Língua Portuguesa.
- Ferraz, L. and Valkhoff, M. F. 1975. "A comparative Study of São-tomense and Cabo-verdiano Creole". In M. F. Valkhoff (ed.), 15-39.
- Ferreira, M. 1985. *A Aventura Crioula*. Lisboa: Plátano Editora.
- Fishman, J. 1991. *Reversing Language Shift*. Clevedon: Multilingual Matters.
- Fleischmann, U. 1984. "Sprache und Politik auf den Kapverdischen Inseln". *Neue Romania* 1: 115-140.
- Hancock, I. F. 1971. A map and list of pidgin and creole languages. In *Pidginization and creolization of languages*, D. Hymes (ed.), 509-523. Cambridge: Cambridge University Press.
- Hancock, I. F. 1986. "The Domestic Hypothesis, Diffusion and Componentiality: An Account of Atlantic Anglophone Creole Origins". In P. Muysken and N. Smith (eds), 71-102.
- Holm, J. 1986. "Substrate Diffusion". In P. Muysken and N. Smith (eds), 259-278.
- Holzer, P. 1991. "Der Stand der Kodifizierung des kapverdischen Kreolisch". In *Zum Stand der Kodifizierung romanischer Kleinsprachen*, W. Dahmen, O. Gsell, G. Holtus, J. Kramer, M. Metzeltin and O. Winkelmann (eds), 313-327. Tübingen: Gunter Narr.
- Kloss, H. 1967. "'Abstand'-Languages and 'Ausbau'-Languages". *Anthropological Linguistics* 9: 29-41.
- Lang, G. 1994. "Ladinização and Capeverdean Creole genesis". Paper presented at the Joint Conference of the SPCL and the SCL, University of Guyana, Georgetown, Guyana.
- Lang, J. 1991. "Die Kategorie Numerus im kapverdischen Kreol". *Neue Romania* 10: 1-19.
- Lang, J. 1993. Das Verbsystem des kapverdischen Kreols (Variante von Santiago). In *Portugiesisch-basierte Kreolsprachen. Akten des 2. Gemeinsamen Kolloquiums der Deutschsprachigen Lusitanistik und Katalanistik*, M. Perl, A. Schönberger and P. Thiele (eds), 137-166. Frankfurt/Main: Verlag Teo Ferrer de Mesquita/Domus Editoria.
- Lopes, E. C. 1967 [1941]. "Dialectos Crioulos e Etnografia Crioula". In Morais-Barbosa (ed.), 405-430.
- Manzini, R. M. and Wexler, K. 1987. "Parameters, binding theory, and learnability". *Linguistic Inquiry* 18: 413-444.
- Meintel, D. 1975. "The Creole dialect of the Island of Brava". In M. F. Valkhoff (ed.), 205-256.
- Mello, H. n.d. "Cape Verdean Creole Portuguese". In *Comparative Creole Syntax*, J. Holm (ed.). MS

- Morais-Barbosa, J. (ed.) 1967. *Estudos Linguísticos Crioulos. Reedição de artigos publicados no Boletim da Sociedade de Geografia de Lisboa*. Lisboa: Academia Internacional da Cultura Portuguesa.
- Morais-Barbosa, J. 1975. "Cape Verde, Guinea-Bissau and São Tomé and Príncipe: The linguistic situation". In M. F. Valkhoff (ed.), 133-151.
- Muysken, P. and Smith, N. (eds) 1986. *Substrata versus Universals in Creole Genesis*. Amsterdam: John Benjamins.
- Nunes, A. 1991. "Quantos falam português". *Nortisul* 2: 18-19.
- PAIGC. 1974. *História. A Guiné e as Ilhas de Cabo Verde*. Paris: Paul Dupont.
- Parsons, E.C. 1923. *Folk-lore from the Cape Verdean Islands* (2 volumes). Cambridge, Mass. and New York: American Folklore Society.
- Pereira, D. 1991. "Crioulos de base portuguesa". *Nortisul* 1: 19-22.
- Peres, D. 1983. *História dos Descobrimentos Portugueses*. Porto: Vertente.
- Perl, M. 1982. *Die Bedeutung des Kreolenportugiesischen für die Herausbildung der Kreolensprachen in der Karibik (unter besonderer Berücksichtigung der kubanischen "habla bozal")*. Dissertation B thesis, University of Leipzig.
- Pires, J. and Hutchinson, J. P. 1983. *Dicionariu Preliminariu Kriolu. Preliminary Creole Dictionary. Cape Verdean/English*. Boston: Funkul Ño Lobu.
- Pires, J. and Hutchinson, J. P. 1994. *Dicionariu Preliminariu Kriolu. Preliminary Creole Dictionary. Cape Verdean/English*. West Newbury, MA: Mother Tongue Editions.
- Römer, R. G. 1991. *Studies in Papiamentu Tonology*. Amsterdam: AWIC.
- Rougé, J. L. 1994. "A propos de la formation des créoles du Cap Vert et de Guinée". *Papia* 3: 137-149.
- Santos, R. 1979. "Comparaison entre le créole du Cap Vert et les langues africaines". In *Seminário sobre a Problemática do Estudo e da Utilização do Crioulo*, 1-28.
- Schuchardt, H. 1888. "Kreolische Studien VII: Ueber das Negerportugiesische von Annobom". *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien* 116: 193-226.
- Schwegler, A. 1992. "Future and conditional in Palenquero". *Journal of Pidgin and Creole Languages* 7: 223-59.
- Scotti-Rosin, M. 1983. "Wege und Ziele bei der Erforschung und Beschreibung kreolischer Sprachen". *Aufsätze zur Portugiesischen Kulturgeschichte* 18: 190-201.
- Silva, B. L. da. 1984 [1957]. *O Dialecto Crioulo de Cabo Verde*. Lisboa: Imprensa Nacional, Casa de Moeda.
- Silva, T.V. da. 1988. *Natal y Kontus*. Praia: Institutu Kauverdianu di Libru.
- Silva, T.V. da. 1992. *Tenpu di Tenpu (1168 dibiña tradicional)* Vol. I. Praia: Instituto Caboverdiano do Livro e do Disco.
- Silva, J.V. da. 1994. "O ensino bilingüe em Cabo Verde". *Papia* 3: 108-111.
- Singler, J. V. 1990. "On the use of sociohistorical criteria in the comparison of creoles". *Linguistics* 28: 645-659.
- Soares, A. M. 1947. *O Dialecto Crioulo de S. Vicente de Cabo Verde. Contribuição para um Estudo do Crioulo Caboverdeano*. Coimbra: Faculdade de Letras.
- Suzuki, M. 1994. *The Markers in Cape Verdean CP*. MS.
- Thiele, P. 1991a. *Kabuverdianu: Elementaria seiner TMA-Morphosyntax im Luso-kreolischen Vergleich*. Bochum: Brockmeyer.

- Thiele, P. 1991b. "Zur Numerusdiskussion im Nominalbereich portugiesisch-lexifizierter Kreol-Sprachen". In *Kontakt und Simplifikation: Beiträge zum 6. Essener Kolloquium über "Kontakt und Simplifikation" vom 18.-19.11.1989 an der Universität Essen*, N. Boretzky, W. Enninger and T. Stolz (eds), 175-192. Bochum: Brockmeyer.
- Valkhoff, M.F. 1975 (ed.) *Miscelânea Luso-africana* Lisboa: Junta de Investigações Científicas do Ultramar.
- Veiga, M. 1982. *Diskrison Strutural de Lingua Kabuverdianu*. Praia: Institutu Kabuverdianu di Livru.
- Veiga, M. 1986. *Oju d'Agü*. Praia: Instituto Caboverdiano do Livro.
- Veiga, M. 1995. *O Crioulo de Cabo Verde. Introdução à Gramática*. Praia: Instituto Caboverdiano do Livro e do Disco/Instituto Nacional da Cultura.
- Winford, D. 1997. "Creole formation in the context of contact linguistics". *Journal of Pidgin and Creole Languages* 12: 131-151.

“High” Kwéyòl: The Emergence of a Formal Creole Register in St. Lucia

Paul B. Garrett*

1. Introduction

In St. Lucia, a French-lexified creole known as Kwéyòl co-exists with the country’s official language, English. In the current postcolonial context, this is giving rise to a somewhat unusual result: the emergence of a “high” or formal register of the basilectal creole as an alternative to the official standard language. In most creole situations, the creole and the standard are lexically related; the creole tends to be restricted to informal contexts, while the standard lexifier language predominates in formal and “high-prestige” settings. In such settings, intermediate between the two there is typically a range of mesolectal varieties, such that there is a more or less gradual transition from basilect to acrolect along a “creole continuum.”

But in St. Lucia, Kwéyòl is a language wholly distinct both lexically and grammatically from English, and St. Lucians often characterize their society as “bilingual.” In recent years Kwéyòl has come to be valorized by some as a central feature of post-Independence national culture, worthy of use in the most formal and prestigious settings. As a result of these and related sociocultural and ideological factors, a “high” register of Kwéyòl is now emerging as an alternative to English. This “high” Kwéyòl is being used increasingly on the radio, in advertisements, in public speaking, and in other such formal contexts that formerly were more or less the exclusive domain of standard English.

Although it is used in these contexts as a marked alternative to English, high Kwéyòl can be shown to be the product of anglicization in many respects. (“Anglicization” is used here as a cover term for various processes of language

change attributable to Kwéyòl's sustained contact with English, of which several will be examined below; it is not intended to convey or imply a particular judgment or evaluation of these processes.) Over recent years, Kwéyòl has been demonstrably affected by contact with English at virtually all levels of analysis, and this is especially evident in the high register. Furthermore, use of this register — mostly by a small cadre of well-educated speakers who, unlike most St. Lucians, also command standard English — tends to be regulated by stylistic and generic conventions that have been transferred directly from standard English. The emergence of this high register can thus be shown to be part of a more general phenomenon of anglicization that is affecting Kwéyòl as spoken at all levels of St. Lucian society. Somewhat paradoxically, then, the emergence of the high register, which is the result of efforts to “preserve” and “instrumentalize” Kwéyòl (and which may indeed be a factor favoring its survival at present), is at the same time contributing to the ongoing anglicization of the language overall. This chapter describes the sociolinguistic and ideological setting in which the high register is emerging, and examines the characteristic features that distinguish it from ordinary Kwéyòl.

2. Creoles and the concept of register

High Kwéyòl is referred to here as a *register*, a term that has rarely been used in the study of creole languages. This is perhaps regrettable, since the concept lends itself readily to analysis of some of the central issues that concern creolists. A register is defined in terms of the domains, settings, and contexts in which it is used. These are recurrent and specifically identifiable, and tend to give rise to the development of specifiable markers of various kinds (lexical, phonological, morphological, syntactic, discursive, etc.). In many respects, registers are comparable to what some sociolinguists and others since the time of Labov's early work have referred to as styles. The difference between registers and styles is primarily one of degree, rather than of kind; the main difference, insofar as one exists, is that styles generally exhibit somewhat less structural elaboration than registers. At the opposite extreme (in terms of structural elaboration) from styles are the highly differentiated “high” and “low” varieties associated with cases of diglossia; these can also be regarded as registers (Hudson 1994). What unifies varied applications of the concept of register is that it is aspects of the setting or context in which

language is being used — factors such as social distance of speaker from interlocutor, political or ideological orientation of the addressee(s), differences of age, status, or gender, relative formality of the occasion, etc. — that condition the manner of speaking and the types of linguistic forms used. This is also what crucially distinguishes *register* from *dialect* and other terms that are defined primarily in terms of groups of speakers and/or those speakers’ social characteristics (Ferguson 1994).

Winford (1991:575) notes that such situational variation “has received far less attention than social differentiation” in sociolinguistic studies as a whole, including studies of Caribbean creoles. Among the latter, some notable exceptions that he mentions (aside from various works in the Labovian interview-based tradition) are Abrahams’s (1983) collected works and Escure’s (1982) work on Belize. An earlier example not mentioned by Winford, but particularly relevant here, is Voorhoeve’s (1971) intriguing (if cursory) article on Surinamese “Church Creole” (a formal register of Sranan) and “pagan cult languages.” To these can now also be added Youssef’s (1991a, 1991b, 1993, 1996) work on children’s acquisition of “varilingual” competence in Trinidad, and Rickford & McNair-Knox’s (1994) study of style-shifting by a speaker of African-American Vernacular English. Although none of these researchers uses the term *register*, each does make use of some analogous or partially overlapping concept such as “stylistic variation,” “genre,” or “speech event.” In a still more recent article, Patrick (1997:42) concurs with Winford’s above-cited observation by noting, “Notions of style and register [...] have barely been explored in pidgin and creole linguistics.” Patrick proceeds to analyze two examples of what he calls “internal varieties” of Jamaican Patwa, designating “Speaky-Spoky” a style and “Rasta Talk” a register. He notes, however, “In treating Speaky-Spoky as a style here, and Rasta Talk as a register [...] I do not mean to imply that sociolinguists have developed clear and consistent definitions or robust concepts that distinguish these two — quite the contrary” (p. 44).

Clearly, then, the definition of register is not firmly agreed upon; different researchers have used the concept in somewhat different ways and to different ends. What is certain is that intraspeaker or situational variation of the kinds that the concept of register has been used to address can be observed in creole and non-creole societies all over the world. In this respect the flexibility of the concept as heretofore applied may actually be an advantage. As part of a foundation for the investigation of intraspeaker variation and language-internal stylistic differentiation in general, the concept of register

can serve creolists well while at the same time putting them on conceptual and terminological common ground with investigators of non-creole languages. The concept of register will be used here in examining a case of language change in St. Lucia, in which the introduction of a creole language into new domains is giving rise to a set of distinctive features and characteristics associated with its use in these contexts.

3. Informal/Formal as a basic register distinction in creole settings

The creole languages of the Caribbean are generally associated with domains and contexts of use that can be characterized broadly as vernacular — that is, everyday, colloquial, informal. Creoles tend to be confined to one side of a fundamental dichotomy, deeply rooted in the region's colonial history, that finds expression (in the work of social scientists as well as in local daily life) through various oppositions such as traditional/modern, uneducated/educated, oral/literate, non-standard/standard, African/European, black/white, rural/urban, private/public. Most of these oppositions are applied, in both popular and academic discourse, to many topics other than language; but language usually enters into the discussion at some point. When it does, the creole is set in contrast, implicitly or explicitly, to a “standard” European language, usually (but not always) the standard variety of the language that has provided the bulk of the creole's lexicon.

Although the various oppositions just mentioned are difficult to subsume within a single overarching dichotomy, they do have something in common and can all be regarded as interrelated. With specific regard to language, one possible way of characterizing their inherent commonality is by means of the pairing informal/formal. This pairing (like all of those mentioned above) is actually better conceptualized as the endpoints of a continuum than as a simple bipolar opposition. The continuum is of course a well-established concept in creole linguistics (e.g. Bickerton 1975, Rickford 1987); although it has been critiqued in recent years as representationally and/or theoretically inadequate (Carrington 1992, Le Page & Tabouret-Keller 1985), the continuum remains a useful and parsimonious model for representing those situations where clear-cut cases of societal bilingualism or diglossia do not obtain. Likewise, informal/formal as used here should be regarded not as a pairing of polar opposites, but rather as a gradient dimension of variation.¹

In societies in which the creole co-exists with its standard lexifier language (e.g. Jamaica, Antigua, Martinique), formal speech is generally acrolectal speech. In these situations, the pairing informal/formal is virtually the equivalent of the pairing creole/standard. It is well known that particular creole features, and/or outright code-switching, may be used strategically by skillful speakers in formal settings to lend a flavor of local authenticity to their words, to evoke a sense of solidarity with their audience, etc. Even so, the separation by domain between basilect and acrolect in creole societies is striking and has often been remarked upon. Devonish (1991:585), for example, goes so far as to characterize the Guyanese situation as "diglossic" in nature, based on the general observation, "Varieties approximating to standard Guyanese English tend to be employed in the more public and formal situations of interaction. This contrasts with the more private and informal situations in which varieties approximating to Guyanese Creole tend to be employed." As Devonish goes on to acknowledge, however, Guyana does present a continuum situation (the most extensively studied one at that), and speakers "normally control more than one language variety," using them differentially according to "the social factors present within the speech event and the social functions associated with each of these varieties."

Situations in which the society's designated official (i.e. standard, European) language is not lexically related to the creole, on the other hand, present a special case. Whether or not the creole continuum model applies in these situations has been something of an open question, if one rarely dealt with. Robertson (1982) makes the case that the continuum does apply, quite unproblematically, to the case of Berbice Dutch and English in Guyana, and by implication to other similar cases. But it seems clear that in any such case, some minor modification of the model is called for in order to acknowledge the lexical disjuncture between basilect and acrolect, which is usually (if not always) less gradual than the grammatical shift. Healy (1993) deals with this disjuncture by the use of two "parallel continua" for Sranan and Dutch in Suriname, asserting that two separate but parallel continua are necessary "since there is a definite structural break between SR [Sranan] and SD [Surinamese Dutch] lects." For St. Lucia, Isaac (1986:32-34) proposes a modified continuum-type model which she divides into four "linguistic codes," labelling them "St. Lucian French Creole," "St. Lucian Basilect," "St. Lucian Creolized English," and "St. Lucian Standard English." (Overlapping shadings and broken boundaries are used in her diagram to indicate "that the divisions are not rigid.") The latter three

codes are subsumed within the overarching category “St. Lucian English,” and are thus as a group distinguished lexically from Kwéyòl.

Regardless of how it is graphically represented, where there is a lexical disjuncture, however gradual, in the creole continuum, there is also a disjuncture in the informal-to-formal continuum.² To go from informal to formal speech (for most if not all speakers in these communities) is necessarily to code-switch, i.e. to use an altogether different lexicon. What happens, then, if local sociocultural and ideological circumstances begin to motivate particular types of stylistic and/or register differentiation *within* the basilectal creole? What happens when there arises a need or a desire to speak in a formal fashion without resorting to such a marked lexical change or code-switching?

4. The case of St. Lucia

Such a situation can arise in cases where the creole has been cut off by historical circumstance from its lexifier and put into contact with a different official-standard language.³ As a result of the new contact situation, various processes of language change and restructuring may occur. Among these processes, as the St. Lucian case will illustrate, may be internal differentiation of the basilectal creole.

In St. Lucia, two closely interrelated processes of language change are occurring at present. One can be characterized as a process of attrition, and the other as a process of emergence. Undergoing attrition is Kwéyòl, the French-lexified creole that has been spoken in St. Lucia for about three centuries.⁴ Kwéyòl is showing many signs of change due to contact with English, which has been used as an official language since early in the nineteenth century. In and of themselves, these changes are perhaps not necessarily diagnostic of attrition; it could even be argued that they are in some sense enriching the language. But it must be considered that at the same time that it is becoming increasingly anglicized, Kwéyòl is no longer being reproduced across the generations as it once was: many (perhaps most) St. Lucian children today are acquiring as their first language an emergent St. Lucian English-lexicon vernacular. In the absence of a local label for this vernacular (about which more below), it will be referred to here by the acronym VESL, for Vernacular English of St. Lucia. In its most creole-like or basilectal form, this emergent restructured vernacular, despite its English lexicon, has more in common

grammatically with Kwéyòl than with standard English.⁵ Taken together, these processes present a case of language convergence, which in turn is giving way to language shift. As Kwéyòl is becoming “anglicized,” English in St. Lucia is becoming, one might say, “kwéyòlized” — to such an extent that a new language variety, the aforementioned VESL, is replacing Kwéyòl as the most widely and most commonly spoken vernacular in St. Lucia.

In considering the case of language convergence and shift just outlined, the focus here will be primarily on one side of the overall phenomenon: the changes arising in Kwéyòl due to its contact with English, which, while giving rise to a new register, are probably also contributing to the attrition of Kwéyòl as a whole. Doubtless the single most important factor in Kwéyòl’s ongoing attrition is the non-reproduction of the language — the fact that many children today are not acquiring Kwéyòl from older members of their families and communities. Non-reproduction of Kwéyòl is considered in greater depth in Garrett (1996a, 1996b, 1997), which also provide a more detailed overview of the ongoing phenomenon of language convergence and shift outlined above. The primary aim here is to deal with another, lesser but increasingly significant factor affecting Kwéyòl: the changes arising in this French-lexified creole as a result of its contact, within a specific social and ideological context, with English. These changes are occurring as a result of various simultaneous processes, some fairly obvious, others more subtle. Collectively, these changes are affecting Kwéyòl at virtually all levels of analysis, from the phonological to the typological.

In order to demonstrate how pervasive these processes of language change are, it will be shown that some of this anglicization is, ironically, a direct result of conscious efforts being made by some St. Lucians to preserve and promote Kwéyòl (and, as part of these efforts, to keep it as “pure” as possible). A small but growing number of St. Lucians have in recent years begun to use Kwéyòl in contexts that were formerly the more or less exclusive domain of English, such as radio broadcasting, advertising, formal debates, and public speaking. The result has been the emergence of a formal or high register of Kwéyòl, as distinct from what will be referred to here as “ordinary” (i.e. informal, everyday conversational) Kwéyòl. Use of this high-register Kwéyòl on the radio and in other formerly English-only domains is sometimes regarded as a sign of success in the ongoing efforts of some St. Lucians to “preserve” Kwéyòl and to elevate its status *vis-à-vis* English. Those who use this high register take pains to speak a particularly “pure” and “authentic”

Kwéyòl, resorting to an English word or phrase only where absolutely necessary. They thereby avoid the extensive use of borrowed and assimilated English words that is commonly heard from speakers of ordinary Kwéyòl (as will be discussed later). But a close examination of this high-register Kwéyòl reveals many other, more subtle changes from ordinary Kwéyòl and from the Kwéyòl of earlier generations. Virtually all of these changes can be attributed to the influence of standard English and its prescriptive norms, stylistic conventions, and discursive genres.

Most of high Kwéyòl's developers and speakers seem to be almost wholly unaware of these subtler changes, and it is partly for this reason that they are interesting and important. Collectively, these changes are bringing about a significant transformation of the language. Broadly speaking, Kwéyòl is being anglicized and, in the process, decreolized. Or perhaps it would be better to say "de-kwéyòlized," since Kwéyòl is not decreolizing in the usual sense: it is changing in the direction of standard English rather than in the direction of its main lexifier, French (which is no longer spoken at all in St. Lucia).

5. Overview of language contact in St. Lucia

A brief historical overview of language contact in St. Lucia will help to put these changes into perspective. The current sociolinguistic situation stems from the island's dual colonial history, which began when St. Lucia was first successfully colonized by the French in the mid-seventeenth century. From the beginning, St. Lucia was an object of great contention and numerous military clashes between France and its arch rival for imperial supremacy in the New World, England. Over the next century and a half it changed hands between the competing powers fourteen times. England finally emerged decisively victorious in 1803, and its possession of the island was formalized by treaty in 1814. But England was evidently more interested in simply controlling the island (for strategic reasons) than in settling it. French colonists were able to stay on even after English dominion had been established, and the French cultural influence continued to predominate in St. Lucian colonial life in following decades. The English language, meanwhile, was mostly restricted to high-level official domains. Not until well into the twentieth century did English begin to become more widespread, mainly due to the increasing availability of formal education.

Following a century and a half as a British colony, and then a short period of associated statehood, St. Lucia became a fully independent member of the Commonwealth in 1979. During the colonial period, Kwéyòl had been subjected to the contempt and general negativity that were heaped upon most other Caribbean creoles. (It is likely that Kwéyòl in St. Lucia, with its French-derived lexicon, was denigrated even more severely than the English-lexified creoles being spoken in other British West Indian colonies during the same period.) The official goal, loudly proclaimed at various times, was to eradicate the language altogether (with schoolmasters and teachers charged with the brunt of the responsibility of eradication). Creole-speakers in St. Lucia, as elsewhere, harbored deeply ambivalent attitudes toward their language. So it is hardly surprising that English was retained as the sole official language at St. Lucia’s independence. Kwéyòl has never been accorded official status, although it is acknowledged where necessary — for example, it can be used in a court of law if one is not proficient in English (in which case an interpreter is provided). Kwéyòl has never been approved as a medium of instruction, although it has never been a secret that teachers of young children, especially in rural areas, sometimes have to use Kwéyòl in order to get a point across. (This is becoming less common, however, since most children are now acquiring more VESL and less Kwéyòl at home.)

Against this historical background of official stances ranging from outright repression at worst to somewhat grudging accommodation at best, considerably more positive attitudes and actions have been taken in regard to Kwéyòl in recent years. Since several years before Independence, Kwéyòl has come to be valorized in some circles, and in the public discourse of some St. Lucians, as a central aspect of national and cultural identity.⁶ Although most of this valorization still occurs in non-official domains, e.g. through the work of certain non-governmental and grassroots organizations, its effects have been considerable and far-reaching. It is now unthinkable, for example, for a teacher to punish a child for speaking Kwéyòl in the schoolyard, as was common practice as recently as a generation ago. Indeed, speaking Kwéyòl (and speaking it well) is now a matter of pride and celebration in certain public contexts, most notably during the annual cultural festival *Jounen Kwéyòl* (Creole Day).

Such changes may well continue to gain momentum in the near future, especially in the wake of the St. Lucia Labour Party’s landslide electoral victory, in May 1997, over the long-entrenched United Workers Party. His-

torically, the UWP (which had been in power for all but three of the past thirty-three years) had been noncommittal in regard to language issues, and never made any significant policy initiatives concerning the status of Kwéyòl. In contrast, SLP candidates occasionally made Kwéyòl and its status an issue during the 1997 campaign (although they generally stopped short of making explicit promises of specific legislative or policy changes). Code-switching between English and Kwéyòl was common in speeches delivered at their flamboyant rallies, which contrasted with the staid, conservative image presented by the UWP.

One year after the Labour Party's ascent to power, the St. Lucian press reported that the new government, at the prime minister's initiative, had amended the Standing Orders (rules of parliamentary proceedings) of the House of Assembly. The Standing Orders had formerly mandated the use of English; in past years, politicians who had attempted to use Kwéyòl in a sustained fashion (i.e. for anything beyond a wry comment or humorous remark) during official House proceedings had been called to order and required to switch to English (Carrington 1984:171-2). The amended section — which the current prime minister reportedly described as part of a “quiet revolution” to make government institutions more accessible and relevant to the people — now reads as follows: “The proceedings and debates of the House, inclusive of the records of such proceedings, shall be in the English language, provided that a Member may offer occasional explanation in Kwéyòl.” Such “occasional explanations” will not be recorded in the official parliamentary record, however; furthermore, according to the press account, “the Speaker will use his discretion to ensure that members use the language occasionally and not for an entire speech” (Anonymous 1998). It remains to be seen whether meaningful changes in Kwéyòl's status in education and other areas that affect the daily lives of ordinary St. Lucians will follow on this early, largely symbolic gesture.

6. Language ideology and language varieties

Although a deep-rooted ambivalence toward Kwéyòl persists in St. Lucia, particularly in official and institutional settings, the language is now celebrated by many St. Lucians (in certain festive and nationalistic contexts, at least) as a constitutive feature of post-Independence national culture. Those

who recognize Kwéyòl in this manner as a full-fledged language in its own right often speak of St. Lucia as a “bilingual” society. The situation is actually much more complex than that, however; over the course of recent decades a continuum-like range of English-lexicon lects has been emerging, along with differentiable “internal varieties” (to use Patrick’s [1997] term) of Kwéyòl. But the notion that St. Lucia is a bilingual society is central to the local ideology of language, which does not make the kinds of distinctions that a linguistic investigator might find it useful to make.

For present purposes, an ideology of language will be defined as a more or less coherent and integrated system of culturally informed beliefs and attitudes concerning language and the nature of its relation to various other aspects of social life (for further discussion, see Woolard & Schieffelin 1994, Woolard 1998). One important aspect of language ideology in St. Lucia concerns how local languages are defined and their boundaries set. Although the English-lexicon vernacular (VESL) that has been mentioned here has occasionally been acknowledged to exist in recent scholarly literature (Alexander 1981, Isaac 1986, St. Juste-Jean 1985, Simmons-McDonald 1994), it is not acknowledged by most St. Lucians. St. Lucians are certainly attentive to the fact that some persons speak English better than others — that is, that some speak more in accordance with pedagogical standards (although these standards are very little known to most). But for the everyday purposes of most St. Lucians, English is English, and no further distinctions need be made. VESL does not have a local name and is rarely remarked upon as a distinct language variety, unlike various English-lexified creoles spoken in other “anglophone” countries of the Caribbean.⁷ The polarity between Kwéyòl (with its “French” lexicon) and “English” is so salient for most St. Lucians that it overrides the less obvious differences, which are mainly grammatical differences (and therefore far less apparent), between vernacular and standard, or basilectal and acrolectal, varieties of “English” as locally spoken. In short, this means that “Kwéyòl” and “English,” with their clearly differing lexicons, are regarded as “the two languages” of St. Lucia, which is often called a “bilingual” society. But this is something like saying that Jamaica or Antigua, where various English-lexified creole varieties co-exist with standard English, are monolingual societies (in which the one language spoken is “English”). It should also be noted here that despite efforts now underway to change negative attitudes toward Kwéyòl, St. Lucians still debate, in casual conversation as well as publicly in the newspapers and on the radio, the question of whether or not Kwéyòl is really “a language.” So it is not

surprising that in most circles VESL is not even acknowledged to exist as yet (or, to look at it differently, that it fades into the very broad local metalinguistic category that is labeled simply “English”).

Other aspects of language ideology in St. Lucia will be addressed further below. For the moment, in order to provide a sense of how the range of

Continuum-type diagram showing relations between St. Lucian language varieties

- **St. Lucian standard English**

Comparable to other national standards, e.g. Australian, American; distinguishable from these others by relatively minor differences of lexicon, phonetics, prosody, and phrase construction.

- **Mesolectal Vernacular English of St. Lucia (VESL)**

Mostly English lexicon with a few Kwéyòl items; some calquing on Kwéyòl constructions; some standard English morphology and use of auxiliary verbs.

- **Basilectal Vernacular English of St. Lucia (VESL)**

Primarily English lexicon, supplemented with some Kwéyòl items; grammatically quite similar to Kwéyòl, with many constructions calqued directly on Kwéyòl constructions.

-----“soft” lexical boundary-----

- **“High” Kwéyòl**

Frequent use of verb phrase anglicisms (calques on standard English phrases); use of verb + satellite anglicisms; occasional “kwéyòlization” of English verbs and other parts of speech; use of neologisms and “false” Kwéyòl words based on English or (less often) French words; use of “authentic” but semi-archaic forms such as the intensifier *twé* and the plural marker *lé*; frequent use of passive constructions (as in formal standard English); avoidance of characteristically creole emphatic and focusing processes such as reduplication, left-dislocation, and predicate clefting; application of standard/literate English stylistic and generic norms; avoidance of the many assimilated English nouns and function words commonly used in “ordinary” anglicized Kwéyòl (cf. below).

- **“Ordinary” anglicized Kwéyòl**

Contains many fully assimilated English lexical items, as well as frequent English borrowings and code-switches; many calques on English/VESL phrases; frequent use of certain English verbs, adverbials, grammatical morphemes and other fully assimilated English items, often to the exclusion of their Kwéyòl equivalents.

- **Kwéyòl as spoken in past generations,**

or as spoken today by elderly monolingual rural St. Lucians.

language varieties currently spoken in St. Lucia might be conceptualized, a continuum-type diagram is provided here. The diagram gives a sense of where high Kwéyòl lies in relation to other discernible St. Lucian language varieties. The characteristics associated with these varieties will be explained more fully in the sections that follow.

The “soft” lexical boundary shown is intended to indicate the area in which the disjuncture between English lexicon and Kwéyòl lexicon occurs; “soft” refers to the fact that some English items will be found below it and some Kwéyòl items above it. This is also, for St. Lucian speakers, the point of disjuncture between “English” and Kwéyòl, and thus the point where the single continuum would be broken into two if one were to take Healy’s (1993) parallel continua approach. It is important to note, however, that the grammatical disjuncture is even less distinct than the lexical one; basilectal VESL and the most anglicized varieties of ordinary Kwéyòl are grammatically similar in many key respects (Garrett 1996a).

7. The “anglicization” of Kwéyòl

Given the history of language contact that was outlined above, it is not surprising that Kwéyòl is becoming anglicized at all levels of St. Lucian society. The most obvious place to see this happening is in the lexicon of ordinary Kwéyòl, which has incorporated numerous words of English origin.⁸ Certain of these, such as the nouns *konpyouta* ‘computer’ and *motoka* ‘motor-car/automobile’, are not surprising, since they are associated with technological developments that did not come about until after English (in one form or another) had become established in St. Lucia. But many ordinary Kwéyòl-speakers today also use certain commonplace English verbs for which Kwéyòl equivalents have always existed, such as those shown below:

Verbs of English origin	Attested Kwéyòl verbs replaced
try	<i>éséyé, débat</i>
feel (in the sense ‘to believe’)	<i>kwè</i>
mean	<i>sinifyé, vlé di</i>
hope	<i>èspéwé, ni èspwa</i>
behave (often used in the imperative <i>Behave kò ou!</i> ‘Behave yourself!’)	<i>wété an koté</i>

These are usually not inflected as English verbs, but remain morphologically invariant like Kwéyòl verbs, and are modified through the use of Kwéyòl tense-mood-aspect particles. Thus a speaker might be heard to say *Sa ou ka try fè èk sa?* ‘What are you trying to do with that?’, in which the verb *try* is modified by the nonpunctual aspect marker *ka*.

Several adverbials, such as *just*, *really*, *too*, and *still* are also used very frequently in ordinary Kwéyòl; and various function words are in extremely common use — such as *because*, *then*, *about*, *so*, and *that* — for which Kwéyòl equivalents have always existed.⁹ Yet the Kwéyòl forms seem to be used now only by the oldest monolingual speakers. Younger speakers may not even recognize the Kwéyòl equivalents, and most have never considered the English origins of the words that they use in their place.¹⁰

Adverbs of English origin	Attested Kwéyòl equivalents
just	<i>annèk, abwézan, talè-a</i> , etc.
really	<i>vwé, vwéman, an véyité</i> , etc.
too	<i>osi</i>
still	<i>toujou</i>
Function words	Attested Kwéyòl equivalents
because	<i>pis(ké)/pas(ké)</i>
then	<i>alòs, answit</i>
about	<i>di, konséné, apipwé</i>
so (conjunction)	<i>alòs</i>
that (complementizer)	<i>ki</i>

It is important to note that assimilation of English items such as these is characteristic of ordinary Kwéyòl, not high Kwéyòl, since speakers of the latter take pains to avoid the use of such obvious English elements. But several other kinds of changes affect both ordinary and high Kwéyòl, such as English-influenced semantic shifts. An example is the Kwéyòl word *ézé*, which evidently derives from the French *aisé* ‘at ease, comfortable’ and formerly had more or less the same meaning. But due to the surface similarity of *ézé* to the English adjective *easy*, the semantic field of Kwéyòl *ézé* has now broadened to comprise the area formerly occupied by the Kwéyòl word *fasil* ‘easy (not difficult)’, a reflex of French *facile*, which is now very rarely heard and is unknown to most speakers of the younger generations.

Verb phrase anglicisms constitute another type of change detectable in

both ordinary and high Kwéyòl. These constructions are the result of direct calquing on English verb phrases. It was mentioned previously that the English-lexicon vernacular now spoken in St. Lucia, VESL, is largely calqued on Kwéyòl; these verb phrase anglicisms make it clear that calquing is occurring in the other direction as well. Both ordinary and high-register speakers use these calques frequently. (I found that even well-educated speakers and pro-Kwéyòl activists do not recognize these as calques on English, and are not readily convinced that they are.) In a few cases Kwéyòl verbs are now being used in new ways patterned on English usage, e.g. *Sa ka gadé bèl* ‘That looks nice’, or *I ka gadé akonmsidéwé i pa la* ‘It looks as if he is not there’. While the English verb *to look* can be used intransitively in the sense ‘to appear’, Kwéyòl *gadé* is ordinarily a strictly transitive verb meaning ‘to look at’, as in *Yo gadé mwen* ‘They looked at me’. The meaning ‘That looks nice’ would ordinarily be rendered in Kwéyòl without use of a verb at all: *Sa bèl*. ‘It looks as if he is not there’ would be rendered with some other verb such as *sanm* ‘to seem’ or *pawèt* ‘to appear’: *I sanm/pawèt i pa la*.

Various other verb constructions patterned on English are common in both ordinary and high Kwéyòl. Some examples are shown below:

English-influenced verb constructions Attested Kwéyòl verbs replaced

<i>gadé pou</i> ‘to look for’	<i>chaché</i>
<i>vini douvan</i> ‘to come forward’	<i>avansé, pwézanaté kò’w</i>
<i>wè ki</i> ‘to see that (something gets done)’	<i>asiyé</i>
<i>pwan plas</i> ‘to take place’	<i>wivé, pwan kou</i>
<i>ki ni pou fè évèk</i> ‘that has to do with’	<i>ki ka konsenné</i>

When a broad sampling of constructions such as these is considered, it begins to become apparent that Kwéyòl, which like French is typologically a “verb-framed” language (to use terminology proposed by Talmy [1991]), is now becoming more of a “satellite-framed” language that, like English, relies heavily on a relatively small set of verb particles such as *up*, *down*, *on*, *off*, *into* and *for* in the construction of verb phrases. In English, one can take a single basic verb, such as *to look*, and change its meaning considerably by using it in conjunction with various different verb “satellites:” *to look for*, *to look at*, *to look around*, etc.; whereas in a “verb-framed” language such as French, various unitary, lexically distinct verbs convey these different meanings: *chercher*, *regarder*, *explorer*, etc. Not surprisingly, Kwéyòl with its French-

derived lexicon is a verb-framed language. But there is now emerging a tendency toward use of verb + satellite constructions that are evidently patterned on English constructions. Taking the example of the basic verb *gadé* ‘to look at’, it becomes clear that verb + satellite is becoming a productive process in Kwéyòl as it is in English: a paradigm of *gadé*-based verb constructions is emerging which directly parallels the English paradigm, as shown below.

Verb + satellite neologisms

Attested Kwéyòl verbs replaced

gadé pou ‘to look for’

chaché

gadé apwé ‘to look after/take care of’

otjipé

gadé andidan ‘to look into (a matter or issue)’

èkzanminé, étidyé, etc.

A clear and very common example of this is the first example given, the construction *gadé pou*. A direct calque on English ‘to look for’, this construction has emerged as a now commonly used synonym of the Kwéyòl verb *chaché* (obviously a reflex of French *chercher*). It is noteworthy that (as mentioned previously) Kwéyòl *gadé*, like French *regarder*, translates into English as ‘to look at’; the prepositional meaning ‘at’ inheres in the monomorphemic verb. But many speakers (especially those of the younger generations) do not hesitate to add on the overt preposition *pou* to form *gadé pou*.

8. Laying the groundwork for “high” Kwéyòl

Although the anglicized high Kwéyòl under examination here is of recent origins and is still emerging, the use of Kwéyòl in formal domains is not without historical precedent. Probably the earliest known instance of a creole language being used for official purposes is cited by Taylor (1968:612), who notes that in the early years of the nineteenth century (*Paris, 17 Brimer, an 10 République francé, yon et indivisible* ‘Paris, 17 Brumaire, year 10 of the French Republic, one and indivisible’), Napoléon Bonaparte and his general Leclerc issued a proclamation “*a tout zabitans Saint-Domingue*” (‘to all inhabitants of St-Domingue’). In St. Lucia a few decades later, in 1845, the newspaper *The Palladium* reported on a wedding at Marquis Estate between Mr. Louis Jean Baptiste and Miss Mary Susan “of the abovenamed Estate:” “The happy bridegroom being a native of Africa, the Rev. Pastor explained the obligations and duties of matrimony in the Creole language, in such a manner

as to be perfectly understood” (Anon. 1942). Without the benefit of other such historical records, we can only speculate as to how usual or unusual the use of Kwéyòl for such a purpose was in this era. But the fact that this wedding was reported on, and this aspect of it mentioned specifically, suggests that it was considered remarkable at the time (as was, no doubt, the event itself — a wedding of “extraordinary splendour” and “admirable order” between two persons who presumably were ex-slaves).

At the same time, it is important to note here that use of Kwéyòl by the clergy has a long and apparently unbroken history in St. Lucia, doubtless dating back to early in the French colonial period.¹¹ Both Catholic and Protestant clergy today, especially those serving rural congregations, use Kwéyòl in church services — typically along with English, in proportions that vary according to the demographic profile of the congregation (especially with regard to age). While ritual texts may or may not be translated into Kwéyòl, Kwéyòl is commonly used for sermons, teachings, and discussion of selected biblical passages (which are generally read aloud in English and translated line by line into Kwéyòl).

Kwéyòl is also commonly used in public speaking by campaigning politicians. This is not a new phenomenon either; it was quite necessary years ago when many St. Lucians were Kwéyòl monolinguals, and this is still true today in rural areas where middle-aged and older persons make up part of the intended audience. But it may well be that Kwéyòl is being used by politicians even more frequently today due to greater public acceptance, and even public expectation. High-ranking government officials, executives of non-governmental organizations, and other such persons of influence are heard regularly on Kwéyòl radio news programs and are sometimes special guests on talk shows. It can probably be said that persons in such positions of influence (up to and including the prime minister) who are not proficient speakers of Kwéyòl, or who disdain to use it publicly, are at a real disadvantage today. Another, related sign of changing times — in regard to both the status of Kwéyòl in St. Lucian society and the declining numbers of St. Lucians who speak it proficiently today — is that help-wanted advertisements in St. Lucian newspapers now sometimes specify “fluency in Kwéyòl” among desired qualifications.

Important shifts in attitude and public sentiment toward Kwéyòl such as these are of fairly recent origin, having begun around the time of St. Lucia’s Independence in 1979. It was during the years leading up to that event that

local intellectuals and cultural activists first began making concerted efforts to change the prevailing negative or ambivalent attitudes toward Kwéyòl, and to preserve the language in the face of the “erosive” effects of English, by promoting it as a central feature of St. Lucian culture and nationhood. One of their major goals was to develop an orthography for the language as a step toward “instrumentalizing” it — i.e. standardizing and “modernizing” it, and thereby developing it into a vehicle of communication suitable for use in all domains and for all purposes in St. Lucian society (including education, government, public service and social development, broadcasting, etc.) (Carrington 1983, 1990). Although the process was fraught with difficulties, an orthography was established (by a committee that included St. Lucians as well as non-St. Lucian expert consultants), and the results were made public in 1983 (Louisy & Turmel-John 1983). The aim of popularizing the orthography and making St. Lucians literate in Kwéyòl was realized only to a very limited extent, however; neither was Kwéyòl accorded official status by the government or accepted into official domains as had been hoped.

Meanwhile, a few inspired professional radio broadcasters had been experimenting with the use of Kwéyòl on the air since several years prior to Independence. Radio broadcasting on a consistent basis in Kwéyòl began in 1971 with a five-minute news program for farmers sponsored by the Department of Agriculture; the first full-length program, an hour long and airing six days a week, commenced in 1974 (Charles 1985). Kwéyòl-language programming has gained considerable momentum since then. At present the number of hours of Kwéyòl broadcasting is probably at an all-time high, averaging a total of more than eight hours daily on the the three local stations combined. Some of these hours consist of call-in programs, which are quite popular and receive contributions from persons of all ages and with varying levels of proficiency in Kwéyòl (some of whom feel compelled, under the circumstances, to excuse themselves when they need to use an English word or phrase to get their point across). All three stations now have Kwéyòl evening programs, which partially overlap; during my fieldwork in 1996-97 there was in fact no local English broadcasting to be found between 7:30 and 8:30 p.m.

A noteworthy outcome of these various efforts has been the increasing use and acceptance of Kwéyòl in public contexts as an alternative to English. The high-register Kwéyòl that has emerged as a result is in certain respects specifically suited to use in these contexts; for example, various terms (mostly neologisms) are used in talking about world politics or the economy that

probably would not be used in any ordinary Kwéyòl conversation (speakers would normally switch to English to talk about these things). Although English still predominates, high Kwéyòl is now heard fairly commonly, and is sometimes expected or even demanded, in domains that formerly were the exclusive preserve of standard English — on the radio and in public speeches by government officials, for example. But somewhat ironically, those seeking to preserve and elevate Kwéyòl by using it in such formal contexts are in fact anglicizing it in many ways — perhaps just as much (if in different ways) as are the less educated speakers and semi-speakers of ordinary Kwéyòl whose frequent English borrowings and other “abuses” they are trying to correct. In some important ways that generally go unnoticed by high-register speakers themselves, they are effectively “de-kwéyòlizing” Kwéyòl by contributing to already ongoing processes of convergence between Kwéyòl and English.

9. Distinguishing features of high Kwéyòl

Some of the most striking characteristics of the high register result from its conformity to the stylistic and generic norms associated with standard English formats, especially literate genres. This can be as simple a matter as, for example, the avoidance of what in standard English pedagogical grammars are called “run-on” sentences. In ordinary Kwéyòl, such run-on sentences are common and are deemed quite acceptable; but they are avoided in high-register Kwéyòl just as they are in “proper” English. These kinds of English-based conventions are most obvious in radio broadcasting. Use of Kwéyòl on the radio has increased to the point that now virtually any type of program presented in English is also presented in Kwéyòl at one time or another. This includes newscasts, weather reports, talk shows and call-in programs, death announcements, and even formal debates. For the most part, the program formats are carbon copies of their English-language counterparts. This is most apparent in the case of newscasts in which wire service world news reports received at the studio in English are translated, directly and word-for-word, into Kwéyòl. This often has rather odd results, such as frequent use of passive and anticausative constructions, which are relatively rare in ordinary Kwéyòl. Yet radio broadcasters tend to use the passive voice frequently even in local news pieces that they have obviously prepared themselves. Furthermore, they and other high-register speakers tend to avoid the ordinary type of Kwéyòl

passive or anticausative construction, which is basically a matter of using the transitive verb intransitively (e.g. *kay ka bati* ‘houses are being built’ [lit. ‘houses are building’]) — and which therefore probably strikes many Kwéyòl-speakers who also speak standard English as incorrect, or at least as rustic and inelegant. They tend to opt instead for periphrastic constructions, preposing to the main verb either the auxiliary verb *twapé*, which in literal active-voice usage means ‘to catch, to get accidentally (as a cold)’, or the alternate auxiliary *touvé* ‘to get, to find, to discover’. (This is comparable to forming the passive in English by use of the auxiliary *to get* — an important difference being that *get*-constructions are more common in casual, vernacular English usage than in more formal styles and registers.) This results in periphrastic passives such as *Yo twapé kwazé pa motoka-a* ‘They got/were crushed by the car’, and periphrastic anticausatives such as *Yo twapé kwazé* ‘They got/were crushed’. While such periphrastic constructions do occur in ordinary Kwéyòl, they are relatively rare — far less common than simple passives such as *Motoka-a kwazé yo* and anticausatives such as *Yo kwazé*.

Also significant is the nearly total absence from high Kwéyòl of certain features that are very common in ordinary Kwéyòl. Some of these differences can be attributed simply to careful speaking; for example, there is very little of the deletion of medial consonants and contraction of words (e.g. *ba’ay* for *bagay* ‘thing’) that are common in casual conversational Kwéyòl. Similarly, the shortened forms that exist of certain verbs, such as *vin* for *vini* and *sòt* for *sòti*, are rarely used in the high register. Especially noteworthy is the absence from high Kwéyòl of certain typically creole processes of emphasis, such as reduplication, left-dislocation, and predicate clefting (as shown below). These processes are all characteristic of, and extremely common in, ordinary Kwéyòl (and it is noteworthy that they are just as common in basilectal VESL). But here again, it seems that such features must strike educated speakers as grammatically incorrect, or at least stylistically unsophisticated.

With further regard to absences, it can be said that given the extent to which various English function words like those mentioned previously are used in ordinary Kwéyòl — *really*, *so*, *that*, *because*, etc. — their absence from high Kwéyòl, and the use of more “authentic” Kwéyòl forms in their place, can itself be regarded as a defining characteristic of the high register.

The use of verb + satellite constructions by speakers of ordinary Kwéyòl was mentioned previously. High-register speakers also use such constructions, but not necessarily the same ones used by ordinary speakers. Skilled

Emphatic processes avoided in high Kwéyòl

Reduplication:

In “ordinary” Kwéyòl: *Nonm-la malad malad malad*

In VESL: *The man sick sick sick*

Non-reduplicative means of intensifying same adjective in “high” Kwéyòl:

Nonm-la malad anpil
Nonm-la twé malad
 ‘The man is very sick.’

Left-dislocation:

In “ordinary” Kwéyòl: *Lakay li i ka alé*
 ‘He is going home’
 [‘home’ emphasized]

In VESL: *At his home he going*

In “high” Kwéyòl: Rarely occurs; emphasis indicated by non-syntactic means (e.g. stress, intonation)

Predicate clefting:

In “ordinary” Kwéyòl: *Achté ou achté bagay sa-a?*
 ‘You bought that thing?’ [i.e. it isn’t handmade?]

In VESL: *Buy you buy that thing?*

In “high” Kwéyòl: Rarely occurs; emphasis indicated by non-syntactic means (e.g. stress, intonation)

speakers tend *not* to use certain of the more obviously English-influenced constructions, such as *gadé pou* ‘to look for’, favoring instead more “authentic” verb forms (in this case, *chaché*). Still, the general tendency toward use of English-style verb + satellite constructions is apparent in the high register. Most noticeably, high-register speakers sometimes add particles onto Kwéyòl verbs that would already adequately convey the intended meaning; this yields formal-sounding (if somewhat redundant) constructions such as *hosé anlè* ‘to raise up’. (Compare this to the use of such two-part verbs in English for rhetorical/oratorical effect, as in *We must raise up the downtrodden*.)

In further regard to notions of “authenticity,” high-register speakers also tend to favor certain conservative French-derived forms that, while doubtless

“authentic,” are now mostly archaic — i.e., unfamiliar to many speakers of ordinary Kwéyòl today. Examples of this include the use of *lé* as the plural definite article (cf. the usual plural construction *sé* [nominal]-*la*)¹² and use of the intensifier *twé* (cf. far more common colloquial intensifiers such as *byen*, *anpil*, *an chay*, *toubonnman*).

Another tendency on the part of high-register speakers — and something else that ordinary speakers usually do not do — is to “kwéyòlize” English words when the need arises, i.e. where no “authentic” or attested Kwéyòl word exists, or where none is known to the speaker. Speakers under pressure to use high Kwéyòl in situations such as live radio interviews sometimes create such words on the spot, basing them either on English or (less often) French. This yields “false” (if phonetically plausible) Kwéyòl words such as *wèlayab* ‘reliable’ (cf. various forms suggested by Mondesir’s [1992] Kwéyòl dictionary, e.g. *fidèl*, *asiwé*, *èkzakt*) and *enpwouvman* ‘improvement’ (cf. *avansman*, *pogwé*). More frequently, speakers create novel “Kwéyòl” verbs by phonetically altering an English verb and tacking on the typical Kwéyòl verb ending *-é*. This yields a wide variety of “kwéyòlized” English-based verbs such as those shown:

“Kwéyòlized” English verbs

diskòsé ‘to discuss’
ilèkté/silèkté ‘to elect/select’
wilaksé ‘to relax’
manajé ‘to manage’
wèkògnizé ‘to recognize’
èkspèkté ‘to expect’
asisté ‘to assist, to help’
administé ‘to administer’

Kwéyòl verbs replaced

(i.e. attested alternatives to the
 English-based neologisms at left)

diskité
chwézi
délasé
diwijé, kondwi
wikonnèt
atann, èspéwé, etc.
endé, pòté soukou
mennajé, gouvenné

While most of these can be regarded as nonce creations by speakers who want to avoid code-switching or English borrowings, some seem to be passing into more general use. A good example is *wilaksé* ‘to relax’, which during my fieldwork in 1996-97 could be heard several times daily in a radio advertisement for the most popular local brand of rum; I also heard it used in conversation from time to time. This makes clear that the high register is not completely discrete and separate from ordinary Kwéyòl: some of these anglicisms and/or

neologisms do filter into common usage (especially among younger speakers, who may speak a highly anglicized variety of ordinary Kwéyòl to begin with). Still, high Kwéyòl tends to pose certain problems of accessibility for speakers of ordinary Kwéyòl, as will be explained in the next section.

Before continuing, some brief samples of high Kwéyòl will help to illustrate a few of the characteristic features described above. The first excerpt, taken from a radio news story, concerns a donation of books to a local school:

Lékol-la wisivwé an bon kantité liv ki yo kay sévi pou établi, ah, bibliyotèk yo, sa sé library yo. Minis ki ni wèkonsabilité pou mobilizasyon nasyonal té ka palé diwan séwémoni-a, noté ki pwéasantasyon-a kay endé pou enkouwajé lé pawan èk lé twavayè èk lé titja andan inichyativ yo pou sa pousé douvan an pwogwanm dé aktivité pou enkouwajé manmay pou li. Minis-la di i apwézan ka éséyé pou touvé liv pou lékol Au Leon pou sa mété yo an plas an bibliyotèk osi, an library.

‘The school received a good quantity of books which they will use to establish, ah, their *bibliyotèk*, that is, their library. [The] minister responsible for national mobilization spoke during the ceremony, [and] noted that the presentation will help to encourage the parents and the workers and the teachers in their initiative to develop [lit. ‘to push ahead’] a program of activities to encourage children to read. The minister said that he is currently trying to find books for [the] Au Leon school, to put them in place in [the] *bibliyotèk* also, in [the] library.’

The newscaster here introduces a novel Kwéyòl term, *bibliyotèk*, clearly deriving it from French *bibliothèque*; he offers a parenthetical English translation both times he uses it, knowing that this word will not be familiar to most listeners. He uses the semi-archaic *lé* as a plural marker (*lé pawan èk lé twavayè èk lé titja*), and certain other words that are unlikely to be familiar to speakers of ordinary Kwéyòl, such as *établi* ‘to establish’ and *inichyativ* ‘initiative’. Other kinds of high-register features occur in the closing of this same newscast:

Mwen ka wimèsyé’w pou kouté. Mwen ka bay envitasyon pou jwenn épi mwen ankò, sa sé lèndi pwochenn, lèndi ki ka vini, pou an lòt pwogwanm; ni an bon jounen, èk an bon finisman simenn.

‘I thank you for listening. I invite you [lit. ‘I am giving invitation’] to join me again, that’s next Monday, this coming Monday, for another program; have a good day, and a good weekend.’

Here the broadcaster uses the word *pwochenn* ‘next’, a word that may have been commonly used in the Kwéyòl of past generations but is unknown to many ordinary speakers today; the speaker is apparently conscious of this, since he immediately follows it with the far more typical *lendi ki ka vini* (lit. ‘Monday that is coming’). Then he bids his listeners *ni an bon jounen* ‘have a good day’, evidently patterning this on the formulaic closing typically used in English-language programs; the verb *ni* ‘to have’ is ordinarily not used in this imperative fashion in Kwéyòl. This being a Friday broadcast, he adds, *èk an bon finisman simenn* ‘and a good weekend’. While ordinary Kwéyòl speakers will doubtless understand this in this particular context, they would themselves simply use the English word *weekend*.

A formal public debate in Kwéyòl that was organized by a community service organization provides examples of high-register speech in a somewhat different setting, as well as some metalinguistic commentary on the part of participants. This debate was explicitly framed as an ordered, formal exercise in which Kwéyòl-speakers, and the Kwéyòl language itself, would be put to the test, in a sense; as one of the judges observed in his remarks, *Lanng Kwéyòl sé an lanng ki pa djè fòmàl* ‘The Kwéyòl language is a language that is not very formal’. But high-Kwéyòl features were in evidence from the very start of the event, as when the master of ceremonies used a periphrastic passive construction (as well as an assimilated English borrowing) in setting out the rules: *Pwen kay twapé tiwé si pyès sé spika-a-yo kwazé pyès sé wég-la* ‘Points will be taken away if any of the speakers- they break any of the rules’. In addition to such high-register syntactic forms, the participants (all well-educated young adults) also produced a considerable variety of novel terms, mostly based on English — such as *wézolvé* ‘to resolve’, *pèfonmé* ‘to perform’, *envolvé* ‘involved’, and *konmité* ‘committed’, among numerous others — prompting the master of ceremonies to comment afterward, *Mwen tann yonndé mo oswé-a ki m’a té jenmen tann avan* ‘I heard a few words this evening that I’d never heard before’.

10. Differential access to high Kwéyòl and its domains

It must be considered here that unlike most speakers of ordinary Kwéyòl, high-register speakers are generally persons who are well educated and have good command of standard English. In addition, they are generally members

of the urban middle to upper classes who probably use English, and an acrolectal English at that, much more than Kwéyòl in their home, community, and professional lives. It is therefore not surprising that standard English influences are abundant and evident in high Kwéyòl. Skilled speakers — radio broadcasters, politicians, religious leaders, and others whose livelihoods now depend at least in part on their ability to speak the high register — are less likely than ordinary speakers to use obviously English lexical items in their Kwéyòl. But their high Kwéyòl shows many of the less obvious English influences that were described above, including many that are not found in ordinary Kwéyòl since they are associated with standard/literate English formats and genres.

The gap between educated and non-educated speakers, and closely related issues of social class, then, are important considerations here. It is evident, for example, that some high-register speakers have some knowledge of French, and draw on that knowledge to fill in lexical gaps in their Kwéyòl.¹³ (Some seem to do this based on a rather exaggerated notion of Kwéyòl’s similarity to French; such notions were also expressed to me by St. Lucians who knew no French at all.) These “gallicized” forms are often used to fill in actual gaps in the Kwéyòl lexicon, but occasionally a high-register speaker resorts to an overtly French form when there is already an equivalent Kwéyòl word in common usage. This suggests that either the speaker does not know the ordinary colloquial form, or is deliberately attempting to impart a more “authentic,” Gallic flavor to the high register. (This can be tied in with a general tendency for representations of “traditional, authentic” St. Lucian culture to emphasize and romanticize the French colonial period in the island’s history.)

Inevitably, some problems of mutual intelligibility arise between ordinary and high Kwéyòl. Just as it does not occur to ordinary Kwéyòl speakers that words such as *really* and *because* are not, strictly speaking, Kwéyòl words (or at least were not as recently as a generation ago), it does not occur to most high-register speakers that constructions such as *pwan aksyon kont* ‘to take action against’, while lexically “pure,” are in fact direct calques on English phrases and therefore might be less than clear to a rural Kwéyòl monolingual (or even to a “bilingual” speaker of ordinary Kwéyòl and basilectal VESL). I found that in the rural community where I did fieldwork, where adults still speak mostly Kwéyòl on an everyday basis, persons I knew were unable to understand certain words and phrases used, for example, in Kwéyòl radio news broadcasts. (This first became apparent to me while I was transcribing

recordings of such broadcasts with local consultants.) Even a young school-teacher — certainly one of the most highly educated members of this community, and someone who regards himself as bilingual — once remarked to me that he often cannot understand some of the words and phrases used in Kwéyòl news programs.

As this would suggest, the various neologisms being created as high Kwéyòl's vocabulary expands also give rise to problems for ordinary speakers. Radio broadcasters intent on keeping their Kwéyòl lexically "pure" sometimes face the dilemma of wanting to use a Kwéyòl term (which may or may not be a neologism) but at the same time knowing that it is not likely to be understood by most listeners, who would simply use the English term even when speaking Kwéyòl. In such cases they often use the Kwéyòl term and then parenthetically give an English translation, as when a broadcaster used the construction *alé asou witwèt* (lit. 'to go on retirement') and immediately followed it by saying "*ében* 'retire'" ('or "retire"'). While this is one means by which high-register neologisms could gradually become more accessible to the general population, there is very little to suggest that this is actually happening. It is, after all, considerably easier (and far more appropriate, by current social norms) for ordinary Kwéyòl-speakers simply to use English borrowings and/or to code-switch during the course of everyday interactions than to use forms such as those heard on the radio. Even the less *recherché* high-register terms, such as *finisman simenn* 'weekend', are rarely heard in ordinary conversation.

Specialized vocabulary is typically a defining feature of a register. In the case of high Kwéyòl, mastery of certain productive processes for creating new vocabulary is also a crucial aspect of high-register usage. Many commonly used Kwéyòl words are quite similar to their English (or French) counterparts, just as French and English words that derive from the same etymological roots are similar, e.g. *wensé* 'to rinse', *èskizé* 'to excuse', *mizik* 'music'. By extending the systematic cognatic correspondences that already exist across the lexicons and phonemic inventories of Kwéyòl and English (and/or French), well-educated high-register speakers can readily come up with a formidable array of sophisticated terms, technical and otherwise, that offer the abstractness, precision (or vagueness), and other characteristics that communication in high-register domains (or concerning topics that would ordinarily be discussed in standard English) often demands — terms such as *twanzaksyon* 'transaction', *watifyé* 'to ratify', *pwémédité* 'to premeditate'. A preponder-

ance of such terms (which the speaker himself or herself may very well never have heard used before in Kwéyòl, and which are certainly unfamiliar to most ordinary speakers) is one of the most immediately apparent features of high-register Kwéyòl.

This suggests that mastery of high Kwéyòl presupposes mastery of standard English (and ideally, some knowledge of French as well). This tends to put high Kwéyòl somewhat beyond the reach of many ordinary St. Lucians, who in addition to not being able to speak it effectively will have some difficulties understanding it — just as they have difficulties when faced with situations in which they need to speak and understand standard English (e.g. when seeking employment today in many sectors of St. Lucia’s developing economy). High Kwéyòl exhibits a high degree of direct mutual translatability with standard English, at the level of individual words as well as at the discursive level. This ready translatability goes hand in hand with its functional equivalence with standard English as a language specifically suited for use in certain domains. A parallel can be seen here to the case of language convergence that Gumperz & Wilson (1971) found in Kupwar, a village on the Indo-Aryan-Dravidian border in India. The three codes spoken in Kupwar, Kannada, Marathi, and Urdu, are grammatically and typologically, as well as lexically, distinct, as defined in normative terms. But as spoken locally in Kupwar after generations of sustained contact, the three codes reveal “a single syntactic surface structure” (p. 155) and “an extraordinary degree of translatability from one local utterance to the other” (p. 154). Clearly the St. Lucian case differs in many respects, but the issue of intertranslatability highlighted by Gumperz & Wilson is an important one here as well. High Kwéyòl’s numerous parallels to standard English — lexical, syntactic, discursive, and otherwise — are not merely coincidental; they serve the practical purposes of those who command both of these “high” language varieties and need to be able to switch readily between them. But this direct equivalence, which may continue to become more pronounced as high Kwéyòl’s vocabulary rapidly expands and becomes conventionalized (if not standardized) within the small cadre of its well-educated speakers, can sometimes have the unintended effect of leaving speakers of ordinary Kwéyòl out of the equation.

11. Instrumentalization and intellectualization as motivations for the development of new registers

Those St. Lucians who wish to promote and preserve Kwéyòl and introduce it into new domains are faced with multiple and sometimes conflicting imperatives. They must develop ways of using Kwéyòl to talk about topics that up until recently have been talked about almost exclusively in English; they must (or at least wish to) keep Kwéyòl as “pure” as possible (lexically, at least) and avoid code-switching and English borrowings, while at the same time finding ways of expanding Kwéyòl’s lexicon in a situation in which contact between the creole and its original lexifier, French, has long since been broken off; and they must modernize, instrumentalize, and for some purposes, “intellectualize” the language, while also striving to keep the developing high register as accessible as possible to ordinary speakers.

The term “intellectualization,” first used by Havránek (1964 [1932]:6), refers to a process of “adaptation to the goal of making possible precise and rigorous, if necessarily abstract, statements, capable of expressing the continuity and complexity of thought;” it affects “primarily the lexical, and in part, the grammatical structure” of a language. Gonzalez (1992:302), following on Havránek’s use of the term, considers intellectualization a process of “cultivation” that leads to “the rise of registers, involving both lexical enrichment through terminology and expanded grammatical uses for dealing with abstract reporting referring to displaced realities.” But how is all of this to be accomplished without what some would consider undue reference to another language, in this case English? Writing about Filipino, Gonzalez (1992:307) sums up the dilemma, which applies to languages in postcolonial settings all over the world:

[T]he expanded use of the developing language into domains hitherto reserved for the colonial language demands the creation of registers (discourse about specific fields) in that language. This creation depends to a large extent on a source language [...] And since scientific academic discourse in the world has been confined mostly to a few western languages and one Asian language in the past two hundred years (for some disciplines, even fewer years), of necessity, the source language for academic scientific discourse in the developing language affects the direction of text-building in an intellectualizing language.

This state of affairs is strikingly reflected in the name of St. Lucia’s leading organization devoted to the preservation and promotion of St. Lucian

creole language and culture, the Folk Research Centre/*Plas Wichès Foklò*. The Kwéyòl rendering of the organization’s name has to cope with two English-language concepts, *research* and *folk(lore)*, for which there are evidently no attested Kwéyòl equivalents. While *foklò* is clearly a kwéyòlized form of the English word *folklore*, *wichès* is a more interesting case. On my first visit to St. Lucia in 1993, the organization’s name was rendered *Plas Wichach Foklò*. (Like *foklò*, *wichach* is simply a kwéyòlization of the English term, or, possibly, of the French *recherche*.) When I returned in 1996, I found that *wichach* had been changed to *wichès* — evidently a play on the attested Kwéyòl word meaning ‘richness’ or ‘wealth’.

“Intellectualization” and “instrumentalization” of a language generally entails standardization. In regard to the “English-lexicon speech communities” of the Caribbean, Devonish (1983) has proposed the establishment of a regional “Institute for Creole Language Standardisation and Development,” the major goal of which would be to develop creoles into “National Official Languages.” In each creolophone society, the resultant standardized creole would become the accepted “Public-Formal language variety” used in the mass media, education, government, etc. — i.e. it would take over the role currently filled by standard English. (Devonish also proposes coordination of efforts at the regional level so that the various national standards would follow “standardised methods” for generating “new, modern technological” terminology [p. 309]. It is not clear whether or not he is also proposing that a single common orthography be established for all of the region’s creoles.) Devonish suggests that such an initiative should begin with radio broadcasting, which could serve as a “catalyst” in the development of “a Public-Formal style of Creole which would hopefully gain gradual public acceptance” (p. 311) — a process already underway in Jamaica (Shields-Brodber 1997) and probably other Caribbean countries as well as in St. Lucia (although it is not clear that the results in these cases have been as Devonish would hope). Elsewhere, Devonish (1991) has outlined a proposal for dealing with the obstacle to standardization posed by continuum-type variation, suggesting that in the Guyanese case language planners could isolate and codify a single variety comprising those features that have the broadest co-occurrence distribution on the continuum (a task already embarked upon by Devonish in his doctoral research).

But standardization — particularly standardization of the sort optimistically proposed by Devonish, in which one language variety among several is

chosen, elevated to a new status, and set forth among speakers of the various other varieties as the “official” language (and thus, inevitably, a language of authority and power) — is always a deeply transformative process, for the language in question as well as for the society in general. This being the case, it also tends to be a highly controversial and politicized process. As Schieffelin & Doucet (1998) show in regard to Haitian, there may arise the thorny question of which dialect or variety is to be anointed as the “authentic” one in the first place, and thus deemed worthy of being standardized and written down. Even if this can be established, the development of an orthographic system for representing the chosen variety (or the selection of one system from among competing systems, as in the Haitian case) can then be a highly sensitive and politically charged arena in its own right, one in which complex social issues that extend far beyond language come into play. The same is true of the various other related processes mentioned here (instrumentalization, modernization, intellectualization), all of which Kwéyòl in St. Lucia, like various other languages in the Caribbean and throughout the world, is currently undergoing.

12. The role of ideology in processes of language change

The ideological dimension of these processes — and closely related issues of power — must not be overlooked. In regard to standardization, Jaffe (1996: 828), citing Eckert (1983) and others, points out an important paradox: the process “can and does create new forms of domination and linguistic alienation by setting up models of correctness that are far removed from local speech.” Similarly, as Hill (1985, 1998) shows, notions of “purity” and “authenticity,” when imposed by a local elite on a language to which such notions formerly did not apply, can even lead to a type of “linguistic terrorism” — one that essentially re-creates, in relation to the local in-group code, the same type of alienation that characterizes non-elite speakers’ relation to the dominant official language. In such situations, as Jaffe points out in regard to Corsican, efforts to preserve and valorize the local language (e.g. by introducing it into domains formerly occupied exclusively by the standard-official language) are, certainly, acts of resistance to domination; but they are acts that, while “challeng[ing] dominance itself,” ultimately fail to “challenge dominant criteria of value.” These latter have typically become such an integral part of local ideologies of language that they are applied in similar if

not identical ways to both of the languages in question.

These “dominant criteria of value” are part and parcel of what can be characterized broadly as a “Western” ideology of language — one that can be traced back to the consolidation of the western European nation-states, influenced the course of imperial expansion and colonialism, and in many respects remains profoundly influential on a global scale today (see e.g. Anderson 1983, Grillo 1989, Blommaert & Verschueren 1998). Particularly relevant here among its central tenets are the notions that a language (that is, anything worthy of being regarded as a language, as opposed to a “dialect,” “patois,” etc.) is coterminous with, and a prerequisite of, national identity (or in some cases, especially today, ethnic identity); that it is a discrete, homogeneous, bounded entity that should be kept “pure” of “foreign” influences (especially lexical incursions); and that it must adhere to certain types of prescriptive norms and standards if it is to be used in certain domains and for certain purposes. As can be seen in the St. Lucian case (as in cases throughout the Caribbean and the world), the profoundly hegemonic nature of these ideas makes them, and their ramifications in practice, almost wholly inaccessible to language planners and cultural activists. Their status as common-sense assumptions among specialists and laypersons alike makes them difficult even to perceive, and more difficult still to bring to public awareness and to challenge in practical terms.

In St. Lucia as elsewhere, language and ideas about language are inextricably intertwined with highly contested issues such as nation-building, economic self-determination, and postcolonial cultural identity. As negotiations over these and related issues unfold in St. Lucia today, Kwéyòl is being transformed in various ways. The emergence of the high register as discussed here is one such transformation. It could be argued that as the high register emerges, Kwéyòl as a whole is in some sense evolving rather than declining. Certainly it should not be overlooked that the sort of on-line creativity that is involved, the ability to draw on existing linguistic resources that comes into play when a speaker draws on English and creates a Kwéyòl-sounding verb on the spot without missing a beat, has been a crucial element in the rise of creole languages in the first place. There can be little doubt that the valorization of Kwéyòl and use of the high register is, for the present historical moment, a factor favoring the language’s survival (albeit in an altered form that continues to change). But while speakers of high Kwéyòl may be preserving (if only among themselves) a few “authentic” forms, such as the plural marker *lé* and

various function words that have fallen out of use among most speakers of ordinary Kwéyòl, they are at the same time anglicizing the language in many other ways that generally go unnoticed. And it is precisely because they go unnoticed that these latter innovations may ultimately prove to be of greater consequence than might be suspected by those persons whose conscious efforts are behind the development of the high register.

Some of these efforts have clearly had positive results, particularly in regard to public perceptions of and attitudes toward Kwéyòl. During my fieldwork, St. Lucians told me that it is now acceptable to use Kwéyòl in virtually any public context, such as in banks or government offices, and that this would have been out of the question not many years ago. I rarely witnessed speakers using Kwéyòl in these settings, however, and many St. Lucians express reservations about doing so — their main concern being that they will be judged as not having sufficient proficiency to use English, and will be stereotyped as backward rustics. Despite the language's increasing valorization in recent years, such negative associations clearly persist for many St. Lucians (or at least, they fear that these associations persist in the minds of others). Similarly, many St. Lucians who express pride in Kwéyòl do not speak it to their children, nor allow their children to speak it, out of concern that the children's developing English will be corrupted (Garrett 1996a, 1996b, 1997).

Clearly, as these disjunctures between idea or attitude and actual practice suggest, Kwéyòl's social-structural relationship to English in St. Lucian society cannot be ignored; it must be fully and realistically taken into account even as efforts to change that relationship are ongoing. It seems clear, for example, that high Kwéyòl will have to draw on English as its lexicon develops. Gonzalez (1992) suggests as much, and Devonish (1983:315), though he acknowledges in his regional-level standardization and development proposal that this may present a problem for St. Lucia in its efforts to integrate with French-official creolophone societies such as Martinique, is able to present only a very tentative and partial solution in passing.¹⁴ For better or worse (if such a judgment can or should be made in these regards), high Kwéyòl will doubtless continue to become more and more anglicized in coming years. At the same time, ordinary Kwéyòl, as a vernacular and medium of everyday communication, will continue to go into decline, as have the French-lexified creoles once widely spoken in other nearby territories where English later became established, such as Trinidad, Grenada, and Carriacou (Parkvall

1997). Indeed, at present it is difficult to imagine how Kwéyòl can be “preserved” in any sense as long as many St. Lucian children are not learning it as a medium of vernacular expression, but are instead acquiring VESL (or more standard varieties of local English, in some cases) as their first language. As Garrett (1997) explains, this too is largely a matter of ideology, one upon which the valorization of Kwéyòl and the elaboration of high Kwéyòl can have only limited effect.

The emergence of high Kwéyòl, like other processes of language change currently underway in St. Lucia, is thus being driven by ideological factors in addition to sociolinguistic and politico-economic factors. To be sure, political and economic factors, at the local level and up to the international level, are crucial; a clear example is the continuing dominance of standard English and its speakers in postcolonial St. Lucia, and the closely related fact that, with few exceptions, only those Kwéyòl-speakers who have had the opportunity to master standard English (and then to acquire other credentials for which standard English is a prerequisite) are able to gain access to the official podiums, broadcasting technology, and other crucial elements of the settings in which high Kwéyòl is emerging. But the case of high Kwéyòl also shows that in efforts to understand and account for processes of language change, local ideologies of language must also be taken into account and thoroughly explored. These ideologies underlie speakers’ attitudes toward languages, their ideas about the differing natures of different languages (or varieties thereof), and their beliefs about what can and cannot, should and should not, be done with and done to particular languages or language varieties. These are important factors in language change that tend to be overlooked if one relies too heavily on, or fails to problematize, established explanatory concepts such as decreolization, relexification, and language shift. In any situation of language change, one has to consider not only what speakers are doing and saying with their language(s), but also what they *think* they are doing and saying — and why they think what they think.

Notes

- * Field research in St. Lucia (January 1996–April 1997) was funded by a Fulbright Fellowship, a National Science Foundation Dissertation Research Grant (#SBR-9522567), and a Wenner-Gren Foundation for Anthropological Research Predoctoral Grant (for the final phase of fieldwork). I am grateful to these three organizations for their

support, as well as to the Spencer Foundation for post-fieldwork support in the form of a Spencer Dissertation Fellowship for Research Related to Education. I thank Bambi B. Schieffelin, John V. Singler, and Constance R. Sutton for guidance; and John H. McWhorter and an anonymous reviewer for helpful comments on a preliminary draft.

1. “Prestige” is a related concept that has been favored by some sociolinguists. It tends quickly to become problematic, however — due, for example, to the need to take issues of “covert” prestige into account, as pointed out by Reisman (1970) and various others.
2. The lexical disjuncture may in fact be fairly abrupt, and tends to occur close to the basilectal end of the continuum. This is largely due to the fact that strongly normative notions of correctness and purity such as are applied to standard European languages generally do not exist for creoles (which in some cases are not regarded as true “languages” even by their speakers). While a great deal of lexical (and other) acrolectal influence on the basilectal creole is likely to be tolerated and may well even be positively evaluated, the reverse is usually not the case.
3. In this regard it is noteworthy that the Caribbean creole that has thus far been most successfully standardized and “instrumentalized” (i.e. instituted as a language of literacy, instruction, official communication, etc.) is Papiamentu — a primarily Spanish-lexified creole that co-exists with Dutch.
4. St. Lucia’s French-lexified creole will be referred to here as “Kwéyòl,” but it should be noted that this name for the language is at present used by few St. Lucians — mainly by certain intellectuals, cultural activists, radio broadcasters, and others who are striving to promote and preserve the St. Lucian creole language and traditional culture. Persons in the community where I conducted field research, and indeed the great majority of St. Lucians, ordinarily call the language “Patwa.” This is by no means necessarily a derogatory label; but some persons who are conscious of its colonial origins prefer to call the language “Kwéyòl” instead. I have chosen to use “Kwéyòl” here because I believe that this label is gradually passing into more general use.
5. Several researchers (e.g. Isaac 1986, Simmons-McDonald 1994) have acknowledged in some fashion the existence of a St. Lucian English-lexicon vernacular. The term “creole” has often been used in designating it (e.g. “St. Lucian Creole English”), but this is misleading. Although it is a product of restructuring through contact with a creole language (Kwéyòl), the language variety referred to here as VESL has not arisen through processes of “creolization” in the sense that term is ordinarily understood; I therefore prefer to call it instead a “vernacular” (Garrett 1998).
6. The persons and organizations involved in these efforts have also been behind the movement to stop calling the language “Patwa” and to call it “Kwéyòl” instead. Aside from concerns about the colonial origins of the label “Patwa” (or “patois”), this arose primarily out of their desire to foster a sense of a distinct St. Lucian creole culture, and out of a conviction that the creole language is an essential part thereof.
7. St. Lucians occasionally speak of “broken English,” but this does not refer to a particular non-standard vernacular variety — it refers to any type of “English” that is quite literally “broken,” i.e. interspersed with Kwéyòl words and phrases (code-switches). While “broken English” is negatively evaluated (and often ridiculed), Kwéyòl interspersed with English borrowings and code-switches, even heavily so, is not — except perhaps in those contexts in which high Kwéyòl (which ideally is lexically “pure”) is called for.

8. Allen (1994a) examines this issue, considering the pros and cons of four different potential ways of conceptualizing it (relexification, decreolization, recreolization, and adlexification), and concludes that the notion of “adlexification” is most fitting.
9. These various words of English origin have been assimilated phonologically as well as grammatically into Kwéyòl. Thus for example the word *just* as used by Kwéyòl speakers would actually be better represented in Kwéyòl orthography as *djòs*. In the interest of clarity, however, all such words are shown here in English orthography. Also noteworthy in this regard is that the assimilation of English lexemes such as *try* and *fridj* ‘refrigerator’ has resulted in the introduction of a new phoneme, /r/, into Kwéyòl; this matter is considered in some depth by Allen (1994b).
10. It might be argued that these English elements are being used for pragmatic effect, and should therefore be regarded as pragmatically motivated code-switches, but in most cases they are not regarded as such by the speakers themselves. Furthermore, these elements generally reflect true gaps, as opposed to momentary lapses, in the Kwéyòl vocabularies of young to middle-aged speakers.
11. Strobel-Köhl (1994:48-49) cites a catechism written in Kwéyòl (using French orthographic conventions) which is attributed to a Fr. Claustre and dates from c. 1900. While this is noteworthy as an early example of written Kwéyòl, spoken Kwéyòl was doubtless being used for catechism and similar purposes long before this.
12. In regard to the use of *lé* as plural marker, Carrington (1984:67fn) notes, “The feature is sporadic, apparently sophisticated but neither predictable or [sic] productive. It is limited to a small number of speakers although most persons understand it, and it occurs mainly in antiquated or traditional expressions.”
13. Regarding the increasing use of Kwéyòl in radio broadcasting, Dalphinis (1985:52) observes: “This new wave of enthusiasm [...] is, however, resulting in the use of a specialised Patwa [...] This is as a result of the use of French items taken from French dictionaries by broadcasters from Castries whose knowledge of Patwa lexicon was [sic] at times relatively limited, and the need to make use of specialised vocabulary to describe items and events formerly outside the traditional domain of Patwa, for example, world nuclear disarmament.”
14. Carrington (1990:79) suggests in passing that a “modernized” Kwéyòl could draw on “other varieties of Antillean which, by virtue of differences in their environment, have developed new vocabulary that can be applied to St. Lucian needs.” Although he does not delve into the practical difficulties that this would entail, he does acknowledge the need for “careful attention to the public acceptability” of such innovations.

References

- Abrahams, R.D. 1983. *The Man-of-Words in the West Indies: Performance and the Emergence of Creole Culture*. Baltimore: The Johns Hopkins University Press.
- Alexander, L. 1981. “Patois influence in St. Lucian English speech”. Typescript. University of the West Indies at Cave Hill: Caribbean Studies Project.

- Allen, J. 1994a. *Sainte-Lucie: Relexification, Décréolisation, Recréolisation, ou Adlexification?* Unpublished *mémoire de D.E.A. des sciences du langage*, Université Lumière Lyon II.
- Allen, J. 1994b. "Can English loanwords create new phonemes in St. Lucian French Creole?" Paper presented at the Workshop on Developmental Issues in Creole Languages and Linguistics, University of Westminster, United Kingdom.
- Anderson, B. 1983. *Imagined Communities: Reflections on the Origins and Spread of Nationalism*. London: Verso.
- Anonymous. 1942. "St. Lucia 97 Years Ago" (an extract reprinted from *The Saint Lucia Palladium* of 10 May 1845). *The Voice of St. Lucia*, 10 December 10.
- Anonymous. 1998. "Kwéyòl in the House". *The Saint Lucia Mirror*, May 22, p. 3.
- Biber, D. and Finegan, E. 1994 (eds). *Sociolinguistic Perspectives on Register*. New York: Oxford University Press.
- Bickerton, D. 1975. *Dynamics of a Creole System*. Cambridge: Cambridge University Press.
- Blommaert, J and Verschueren, J. 1998. "The role of language in European nationalist ideologies". In Schieffelin, Woolard and Kroskrity, 189-210.
- Carrington, L.D. 1983. "Preparing a creole for a developmental task: The case of Antillean French-lexicon Creole". *Carib* 3:69-80.
- Carrington, L.D. 1984. *St. Lucian Creole: A Descriptive Analysis of its Phonology and Morpho-Syntax*. Hamburg: Helmut Buske.
- Carrington, L.D. 1990. "The instrumentalization of St. Lucian". *International Journal of the Sociology of Language* 85:71-80.
- Carrington, L.D. 1992. "Images of creole space". *The Journal of Pidgin and Creole Studies* 7(1):93-99.
- Charles, E. 1985. "The development of Kwéyòl usage in radio in Dominica and St. Lucia". Typescript in the library of the Folk Research Centre/Plas Wichès Foklò, St. Lucia.
- Cheshire, J. (ed.) 1991. *English Around the World: Sociolinguistic Perspectives*. Cambridge: Cambridge University Press.
- Devonish, H. 1983. "Towards the establishment of an institute for creole language standardisation and development in the Caribbean". In *Studies in Caribbean Language*, L.D. Carrington in collaboration with D.R. Craig and R.T. Dandará (eds), 300-316. University of the West Indies, St. Augustine: Society for Caribbean Linguistics.
- Devonish, H. 1991. "Standardisation in a creole continuum situation: The Guyana case". In Cheshire (ed.), 585-594.
- Eckert, P. 1983. "The paradox of national language movements". *Journal of Multilingual and Multicultural Development* 4:289-300.
- Escure, G. 1982. "Contrastive patterns of intra-group and inter-group interaction in the creole continuum of Belize". *Language in Society* 11:239-264.
- Ferguson, C.A. 1994. "Dialect, register, and genre: Working assumptions about conventionalization." In Biber and Finegan, 15-30.
- Garrett, P.B. 1996a. "Language socialization and the emergence of an English-lexicon vernacular in St. Lucia". Paper presented at the eleventh biennial conference of the Society for Caribbean Linguistics, St. Maarten, Netherlands Antilles.
- Garrett, P.B. 1996b. "Language socialization and language change in St. Lucia". Paper

- presented at the annual meeting of the American Anthropological Association, San Francisco.
- Garrett, P.B. 1997. "Socializing ambivalence: Variation in language socialization practices and convergence of codes in St. Lucia". Paper presented at the annual meeting of the American Anthropological Association, Washington, D.C.
- Garrett, P.B. 1998. "An 'English Creole' that isn't: On the origins of St. Lucian English-lexicon vernacular". Paper presented at the joint meeting of the Society for Pidgin and Creole Linguistics and the Society for Caribbean Linguistics (in conjunction with the annual meeting of the Linguistic Society of America), New York.
- Gonzalez, A. 1992. "Reconceptualization, translation and the intellectualization of a Third World language: The case of Filipino". In *Sociolinguistics Today: International Perspectives*, K. Bolton and H. Kwok (eds), 300-322. London and New York: Routledge.
- Grillo, R.D. 1989. *Dominant Languages: Language and Hierarchy in Britain and France*. Cambridge: Cambridge University Press.
- Gumperz, J.J. and Wilson, R. 1971. "Convergence and creolization: A case from the Indo-Aryan/Dravidian border". In Hymes (ed.), 151-167.
- Havránek, B. 1964 [1932]. "The functional differentiation of the standard language". In *A Prague School Reader on Esthetics, Literary Structure, and Style*, P.L. Garvin (ed. and trans.), 3-16. Washington, D.C.: Georgetown University Press.
- Healy, M. 1993. "The parallel continuum model for Suriname: A preliminary study". In *Atlantic Meets Pacific: A Global View of Pidginization and Creolization*, F. Byrne and J. Holm (eds), 279-289. Amsterdam: John Benjamins.
- Hill, J.H. 1985. "The grammar of consciousness and the consciousness of grammar". *American Ethnologist* 12:725-737.
- Hill, J.H. 1998. "'Today there is no respect': Nostalgia, 'respect,' and oppositional discourse in Mexicano (Nahuatl) language ideology". In Schieffelin, Woolard and Kroskrity, 68-86.
- Hudson, A. 1994. "Diglossia as a special case of register variation". In Biber and Finegan, 294-314.
- Hymes, D. 1971 (ed.). *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press.
- Isaac, M.F. 1986. *French Creole Interference in the Written English of St. Lucia Secondary School Students*. Unpublished M.Phil. thesis, University of the West Indies at Cave Hill.
- Jaffe, A. 1996. "The second annual Corsican spelling contest: Orthography and ideology". *American Ethnologist* 23:816-835.
- Le Page, R.B. and Tabouret-Keller, A. 1985. *Acts of Identity: Creole-based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.
- Louisy, P. and Turmel-John, P. 1983. *A Handbook for Writing Creole*. Castries: Research St. Lucia Publications.
- Mondesir, J.E. 1992. *Dictionary of St. Lucian Creole*. (L.D. Carrington, ed.) Berlin: Mouton de Gruyter.
- Parkvall, M. 1997. "The rise and fall of French Creole on the Commonwealth Lesser Antilles". Unpublished manuscript, University of Stockholm.
- Patrick, P.L. 1997. "Style and register in Jamaican Patwa". In Schneider, 41-55.
- Reisman, K. 1970. "Cultural and linguistic ambiguity in a West Indian Village". In *Afro-*

- American Anthropology*, N.E. Whitten and Szwed, J.F. (eds), 129-144. New York: The Free Press.
- Rickford, J.R. 1987. *Dimensions of a Creole Continuum: History, Texts, and Linguistic Analysis of Guyanese Creole*. Stanford: Stanford University Press.
- Rickford, J.R. and McNair-Knox, F. 1994. "Addressee- and topic-influenced style shift: A quantitative sociolinguistic study". In Biber and Finegan (eds), 235-276.
- Robertson, I. 1982. "Redefining the post-creole continuum: Evidence from Berbice Dutch". *Amsterdam Creole Studies* 4: 62-78.
- St. Juste-Jean, M.L. 1985. *A Study of the Evidence of Patwa Interference in the Written Work in the 11+ Age Group in Two Rural and Two Urban St. Lucian Schools*. Unpublished B.A. thesis, University of the West Indies at Cave Hill.
- Schieffelin, B.B. and Doucet, R.C. 1998. "The 'real' Haitian Creole: Ideology, metalinguistics, and orthographic choice". In Schieffelin, Woolard and Kroskrity, 285-316.
- Schieffelin, B., K.A. Woolard and Kroskrity, P.V. (eds) 1998. *Language Ideologies: Practice and Theory*. New York: Oxford University Press.
- Schneider, E. (ed.) 1997. *Englishes Around the World: Studies in Honour of Manfred Görlach, Volume 2: Caribbean, Africa, Asia, Australasia*. Amsterdam and Philadelphia: John Benjamins.
- Shields-Brodber, Kathryn, 1997. "Requiem for English in an 'English-speaking' community: The case of Jamaica". In Schneider, 57-67.
- Simmons-McDonald, H. 1994. "Comparative patterns in the acquisition of English negation by native speakers of French Creole and Creole English". *Language Learning* 44:29-74.
- Strobel-Köhl, M. 1994. *Die Diskussion um die "ideale" Orthographie: das Beispiel der Kreolsprachen auf französischer Basis in der Karibik und des Französischen im 16. und 20. Jahrhundert*. Tübingen: Günther Narr.
- Talmy, Leonard, 1991. "Path to realization: A typology of event conflation". *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, L. Sutton, C. Johnson with Ruth Shields (eds), 480-519. Berkeley: Berkeley Linguistics Society.
- Taylor, Douglas, 1968. "New languages for old in the West Indies". In *Readings in the Sociology of Language*, Joshua Fishman (ed.), 607-619. The Hague: Mouton.
- Voorhoeve, J. 1971. "Varieties of creole in Suriname: Church Creole and pagan cult Languages". In Hymes (ed.), 305-315.
- Winford, D. 1991. "The Caribbean." In Cheshire (ed.), 565-584.
- Woolard, K.A. 1998. "Language ideology as a field of inquiry". In Schieffelin, Woolard and Kroskrity, 3-47.
- Woolard, K.A. and Schieffelin, B.B. 1994. "Language ideology." *Annual Review of Anthropology* 23:55-82.
- Youssef, V. 1991a. "Variation as a feature of language acquisition in the Trinidad context". *Language Variation and Change* 3:75-101.
- Youssef, V. 1991b. "The acquisition of varilingual competence". *English World-Wide* 12(1):87-102.
- Youssef, V. 1993. "Children's linguistic choices: Audience design and societal norms". *Language in Society* 22:257-274.

- Youssef, V. 1996. “Varilingualism: The competence underlying codemixing in Trinidad and Tobago”. *The Journal of Pidgin and Creole Languages* 11:1-22.

From Latin to Early Romance: A Case of Partial Creolization?

Stéphane Goyette¹

J'ai été surpris de constater à quel point le grec populaire est demeuré fidèle au grec classique: cette langue a bien du mal à bouger, alors que notre latin littéraire est devenu comme un idiome étranger par rapport au latin tel qu'on le parle. (...) le latin parlé s'est enrichi d'une foule de mots qui sont jugés trop vulgaires pour figurer en latin littéraire, et la grammaire orale s'est désagrégée. Le peuple n'emploie plus guère que le nominatif et l'accusatif, multipliant les prépositions autour de ce dernier cas (...) L'étudiant latin en arrive à écrire une langue artificielle, et il lui faut faire effort pour la déclamer convenablement. (Monteilhet 1984: 159-160)

(I was surprised to find how faithful popular Greek has remained to Classical Greek: this language scarcely seems able to move. By contrast, our own literary Latin has become a foreign tongue compared to Latin as it is spoken (...) Spoken Latin has enriched itself with a host of words deemed too vulgar to appear in literary Latin, and the grammar of the colloquial language has broken down. The masses scarcely use more than the nominative and the accusative, making great use of prepositions with the latter case (...) Latin students end up writing an artificial language, and need to make an effort to recite it adequately. [trans. by the author])

1. Introduction

The thesis of the present paper is that creolization played a role in the genesis of the Romance languages from Latin. Radical though this position may appear, many creolists having clearly stated their opposition to it (e.g. Chaudenson 1995, in a seeming reversal of Chaudenson 1983; Holm 1988), it is in fact scarcely new, and indeed was first explicitly defended by Indo-Europeanist Antoine Meillet, who on the one hand clearly stated that the changes which separated Latin from Romance were similar to those which separate creoles

from their respective lexifier languages, and which moreover were due to similar social factors:

En cessant d'être la langue d'une cité pour s'étendre à un empire, le latin ne pouvait garder ses délicatesses et son originalité (...) les oppositions ainsi marquées étaient trop fines pour être observées par des gens pour qui le latin n'était pas une langue maternelle et qui l'apprenaient avec une certaine grossièreté. *On sait ce que sont devenues les langues européennes chez les esclaves transportés dans les colonies: les divers "créoles", français, espagnol, hollandais, sont des langues où la grammaire est réduite à presque rien et d'où les nuances anciennes ont été supprimées.* (Meillet 1928a: 236; emphasis mine.)

(In ceasing to be the language of a city to become that of an Empire, Latin could not keep its subtle and language-specific traits (...) such oppositions were too subtle to observe for people who did not speak Latin natively and who learned it in a rough-and-ready fashion. We know what became of the European languages among slaves brought to the colonies: the various French, Spanish and Dutch "creoles" are languages where grammar is reduced to almost nothing and from which former nuances have been eliminated. [trans. by the author])

(It should be made clear, incidentally, that the term "grammaire" refers to inflectional morphology.)

On the other hand, Meillet clearly stated elsewhere that the analogy was only partial, i.e. that the creole languages of the Caribbean, when compared to their lexifier languages, had been restructured in far more extensive a fashion than the Romance languages had been when compared to Latin, for social and linguistic reasons:

Les esclaves nègres des anciennes colonies, en acceptant le français ou l'espagnol, ont entièrement transformé ces langues. Ils leur ont donné une prononciation nouvelle; ils en ont simplifié la grammaire. (...) C'est que les parlers des nègres soudanais diffèrent essentiellement du français et que des hommes qui occupent une situation sociale irrémédiablement inférieure ne sentent pas le besoin d'avoir le beau langage de leurs maîtres, qui, de leur côté, ne font rien pour les corriger.

Au contraire les Gaulois, qui parlaient une langue d'un type semblable au latin, et qui ont adopté le latin pour être des citoyens romains et en avoir les droits, se sont assimilés le système complet de la langue latine. (...) à côté d'hommes qui avaient bien appris la langue, il s'en trouvait qui l'avaient acquise imparfaitement." (Meillet 1928b: 122)

(The black slaves of the former colonies, in accepting French or Spanish, have wholly transformed these languages. They changed the pronunciation com-

pletely: they simplified the grammar (...) This is because the languages of African blacks differ fundamentally from French and because men whose social status remains unalterably low do not feel the need to know the proper language of their masters, who, for their part, do nothing to correct them.

In contrast, the Gauls, who spoke a language similar in type to Latin, and who adopted Latin to become Roman citizens and to have the rights citizenship conferred, internalized the full system of the Latin language (...) alongside men who had learned the language well, there were others who had acquired it imperfectly. [trans. by the author]

This accords well with what will be maintained in this paper; namely, that while no recorded form of European Romance could be called a creole, Early Romance was originally born out of the contact between creolized and non-creolized varieties of Latin (and thus could be termed a “creoloid” or “semi-creole”). While many linguists — predominantly Romance scholars — have commented on the possible relevance of creole studies in explaining the emergence of Romance (see Schlieben-Lange 1977 for a useful list of quotations and references), very little of substance has been published on the subject (*pace* Mufwene 1988).

It should immediately be emphasized that this issue — the role creolization may or may not have played — is quite different from that of the role language contact in general may have played in the birth of the various individual Romance languages and dialects, which has been studied quite extensively both with regards to the role of substratal languages (i.e. the influence Gaulish may have had on French) and that of superstratal languages (i.e. the influence Frankish may have had on French).²

While the vast number of languages superseded by Latin, and the many others later in contact with its Romance progeny, have made it natural to examine the role language contact may have played in the emergence and evolution of the Romance languages, it is altogether surprising that, as remarked above, the very possibility that creolization may have played some role in the history of Latin/Romance has been almost completely ignored since Meillet’s day, in spite of the considerable advances in creole linguistics since then. Considering that Latin, originally the language of a single city, became within a few centuries the spoken *lingua franca* of Western Europe,³ eventually coming to replace most of the languages originally spoken over this vast area, and this furthermore in a context of conquest, population displacement, chattel slavery, and unavailability of formal education to all but the elite (Garnsey & Saller 1987), it must be recognized that there is no reason

whatsoever for excluding the possibility that Latin, to some degree at least, was creolized. Indeed, when comparing its external history to that of other European languages, it must be conceded that creolization is far likelier to have played a role in the history of Latin/Romance than in that of any other language of Europe.

Two questions might well be asked at this stage. First, what exactly is meant by “creolized to some degree”? In the preceding paragraph reference was made to the considerable advances in creole studies made since Meillet’s day: surely this warrants a more precise description of the linguistic situation. Second, why is “Romance” being treated as a single, monolithic entity? If creolization is indeed held to have been a pertinent phenomenon in the linguistic history of the various Romance languages, might this not have been a strictly localized, language-specific or even dialect-specific phenomenon?

In answer to the second question, it should be made clear that no position whatsoever need be adopted or believed in dogmatically regarding the exact chronology of the break-up of “proto-Romance” into various forms of speech, distinct from one another as well as from Latin (see Wright 1991 for discussion). Regardless of one’s position regarding this issue, however, it cannot be denied that when they are first attested in writing, Romance languages share a host of traits with one another, setting them off *en bloc* from Latin.⁴ To be sure, divergences are present, and will be pointed out in due course, but these are far outnumbered by the similarities. This, then, is what justifies treating all of these varieties together.

As for why this should be so, this brings us to the answer to the first question. The sociolinguistic scenario assumed to account for the linguistic facts is the following: Early Romance, a creolized variety of Latin, was born in the third century B.C. (Leonard 1978, who himself, it should be noted, does not hold Early Romance to have been creolized in any way). This birth was due to Roman expansion in Central Italy, which led to the large-scale imposition of Latin on non-Latin-speaking (and, in some cases, such as the Etruscans, non-Indo-European-speaking) peoples. This creolized variety stood at one end of a linguistic continuum, Classical Latin being the other end: the two forms of speech were certainly not perceived as distinct languages, even after the fall of the Empire (Wright 1991). While it is quite possible that secondary creolization occurred in the course of its expansion as the dominant language of the Western Roman Empire, there is no need to suppose Latin to have been creolized more than once. (Paradoxically, the fall of the Roman Empire,

triggering some movements of population, may have contributed greatly to the linguistic Romanization of much of Western Europe; cf. Bonjour, Offler & Potter 1952 for such an explanation regarding the Romanization of Northern Italy.)

Further arguments in favor of an early date for the birth of Romance will be presented at the end of section 3. One very strong piece of evidence in favor of this is the early attestation in writing of many changes differentiating Romance from Latin, such as the use of the preposition *ad* as a genitive case marker, or the replacement of nominative by accusative forms: cf. the following two inscriptions, both taken from tombstones (quoted in Elcock 1960: 26):

- (1) *Hic requiiscunt men
bra ad duus fratres (...)*
'Here lie the remains of two brothers' (Classical Latin: *..membra duorum fratrum*, with the genitive plural).
- (2) *Hic quescunt duas matres, duas filias (...)*
'Here lie two mothers, two daughters' (Classical Latin: *..duae matres, duae filiae*: the forms *duas matres, duas filias* are accusative plural forms).

While these two inscriptions have not been precisely dated, it is remarkable that similar forms abound in the Latin inscriptions of Pompeii (dating from the volcanic explosion of 79 A.D.), suggesting a very early birth for Romance.⁵

2. Methodology

The idea that the linguistic processes which gave birth to creole languages may have played a role in the genesis of some languages or language families is not new: creolization has been theorized to have played a role in the birth of, *inter alia*, Marathi (Southworth 1971), Middle English (Bailey & Maroldt 1977) and Modern Arabic (Versteegh 1984).

What all of these studies have in common, however, is a failure to clearly show that creolization is any likelier to have occurred in the languages studied than in any others. From this point of view, Thomason & Kaufman's (1988) criticism of Bailey & Maroldt's analysis of English is quite pertinent. Examining data from other Germanic languages, Thomason & Kaufman demonstrate

that far from being an aberrant or unusual phenomenon, the drift toward a more analytical structure ascribed by Bailey & Maroldt to language contact is the norm in most Germanic languages, and that there is thus no need to resort to creolization as an explanatory tool.⁶

On this basis it is clear that ascertaining whether or not creolization played any sort of role in the genesis of Romance can only be done via a comparative approach, i.e. by comparing the changes which separated Latin from Romance to the changes which occurred during the same period of time in some other language or language family. Such a language would not only have to be typologically similar to Latin; its external history should be sufficiently well-known to preclude the possibility that it could ever itself have been creolized.

Only one language truly fits both requirements: Greek. Not only does it have a written history which goes back further in time than that of Latin, but linguistically it is sufficiently similar to Latin to allow a meaningful comparison to be made. Its external history, moreover, is sufficiently well-known to allow us to say that creolization cannot be said to have played any significant role in its history. The reason for saying this, rather than “cannot have played any role in its history”, is because there does appear to be some evidence in favor of there having been some role played in the evolution of Greek by second-language or even pidginized varieties of the language (Versteegh 1986). Even were one to concede this, however, it would have to be acknowledged that such influence must necessarily have been quite limited in the case of Greek when compared to Latin/Romance. Unlike Greek, Latin expanded outside its original home to become the dominant spoken language of a vast portion of the Roman Empire, thus coming into intimate contact with a great number of other languages, whereas Greek has expanded little beyond its southern Balkanic home,⁷ implying that however pidginization or creolization may have affected its evolution, such influence must have been considerably less profound than in the case of Latin.

The xenophobic attitude Greeks had *vis-à-vis* non-Greeks, which contrasted sharply with that of the Romans (Todd 1994), also makes it unlikely that creolization played any significant role in the evolution of their language.

“But”, the proverbial attentive reader might at this stage point out, “how then could this comparison enlighten us? For it is perfectly natural for languages to diverge from one another: the Romance languages differ from one another, in that they have diverged differently from their original source

language, Latin. For Greek and Latin to have evolved differently is scarcely surprising. Wouldn't you be proceeding no less arbitrarily than those scholars whom you just criticized, inasmuch as you would be bringing to the fore a notion such as creolization to account for the divergences between Latin and Greek, when such a notion is in fact utterly superfluous?"

The answer is that this objection would be perfectly valid if creolization and normal language change did not differ from one another in terms of their observable linguistic effect. If they do, however, and if Latin is found to adhere more closely to "creole-like" than to "normal" language change, then creolization, far from being some arbitrary *deus ex machina* serving to explain that which may be explicable by other means, would be a useful working hypothesis. If, however, Latin is no more "creole-like" than Greek in terms of its linguistic evolution, then the hypothesis falls flat and there is no need to postulate Latin to have been creolized during the course of its history.

This should make it clear that what is being done here is the testing of a hypothesis, in Popperian fashion, rather than the concoction of a web of speculation. In order to guard against the possibility that Greek, in some ways at least, may be an atypical Indo-European language with regards to its evolution, some consideration will be given to the evolution of other branches of Indo-European in the following pages.

For the purposes of this paper, creolization will be held to differ from normal language change in the following fashion: whereas normal language change involves a gradual see-saw-like movement between the creation of analytical and synthetic structures (Schwegler 1990), creolization involves the complete replacement of synthetic structures by analytical ones (cf. most recently, for this typological trait and others which allegedly define a creole, McWhorter 1998). Subsequently, however, as the creole is transformed according to the laws of normal language change, it may create synthetic structures out of originally analytical ones. This means that over time, a creole language of course sheds its distinctive linguistic cast to become increasingly indistinguishable from other, "normal" languages (all the more so if areal pressure from other, neighboring languages is a relevant factor). This is why the stage of Romance to be examined will be that of our earliest texts: if the hypothesis being examined here (that Latin was creolized as it expanded as a spoken language in the Italian Peninsula) is correct, we would expect to find this best reflected, linguistically, in the earliest attested stages of Romance speech.

A possible objection might be that even if one were to discover a greater degree of morphological conservatism in Greek than in Latin, this could well be explicable in the following fashion: since the (predominantly Latin-speaking) Western Roman Empire fell in 479 and the (predominantly Greek-speaking) Eastern Roman (Byzantine) Empire fell nearly a thousand years later in 1452, might not this conservative aspect of Greek be due to a greater degree of social, political and cultural continuity?⁸

Such an objection would certainly be valid. Indeed, such sociopolitical factors have often been invoked to explain differences within Romance itself. Thus, Elcock (1960), among others, ascribes the differences between Northern French, on the one hand — a highly innovative variety of Romance — and Tuscan Italian, on the other — a highly conservative variety of Romance — to the fact that Northern France was a locus of rapid social and political change due chiefly to the various Germanic invasions, while Tuscany remained for most of its history little more than a sleepy backwater.

There exists, however, a good piece of counterevidence to this claim. To this day there exist, in Southern Italy, dialects of Greek, the origin of which goes back to classical times. Located as they are in Western Europe rather than the Balkans, separated from the rest of the Greek-speaking world, these dialects (known as Italic Greek) provide an ideal testing ground to weigh the claim that sociopolitical continuity was the main factor in making Greek a conservative language in relation to Proto-Indo-European. If, in spite of their sociopolitical isolation from the Greek world, the evolution of Italic Greek does not differ in broad outline from that of Balkan Greek, then the conservative cast of Greek cannot be explained away by sociopolitical factors.⁹ This in fact is what we find (Rohlf 1949).

3. Nominal Morphology

The nominal morphology of Latin and that of Greek were highly similar to one another, making comparison of this particular subset of the grammars of the two languages, along with their subsequent evolution, both easy and fruitful.

In both languages case-marking was fusional, most inflectional endings not segmentable into distinct morphemes indicating number, gender and case. In both languages nouns are divided formally into several declension class types, and indeed the classes of Latin and Greek correspond quite closely.

Finally, the various case forms of the two languages are very close to one another in function, i.e. the accusative in Latin is used very much like the accusative in Greek. The chief difference between the two languages lies in the fact that Greek had no case-form corresponding to the Latin ablative (and thus had five distinct case forms, rather than six as in Latin).

Another significant difference between Latin and Greek (true of nominal as well as of verbal morphology) lies in the system of number: whereas Latin morphologically differentiated a singular and a plural, Greek morphologically differentiated a singular, a dual and a plural. Thus, whereas Latin had *dominus/domini* ‘master/masters’ (nominative), Greek had *ho kyrios/tô kyriô/hoi kyrioi*¹⁰ ‘the master/the (two) masters/the (more than two) masters’ (nominative).¹¹ However, the dual was an optional rather than obligatory morphological category in Classical Greek (Ragon 1961) (one could simply use the numeral *dyo* ‘two’ plus the plural, as in Latin and most modern European languages), and thus we will ignore the dual in the following tables and discussion: it disappeared from Greek a few centuries later in any case.

Some points should be clarified about the following tables: because of divergent vocalic evolution between various Romance languages, the vocalic evolution assumed in the “expected Romance” column is that of Sardinian, generally acknowledged to be the most archaic of the Romance languages. Vowel length, although phonemic in Latin, was not indicated in writing; a colon (:) will be placed after a vowel to note length whenever this is deemed pertinent to the facts being discussed. Finally, the reason why modern Greek forms (taken from Joseph 1987) were used (rather than Greek forms contemporary to early Romance, i.e. ninth to twelve century A.D.) was (if we may anticipate the conclusion) to highlight the remarkably archaic aspect Greek morphology has when compared to that of the attested Romance languages.

3.1. Comparisons between Latin and Greek

Our first comparison is between the development of feminine nouns in Latin and Greek:

Table 1. Feminine Latin *a*- and Greek *ê*-stems

CASE	Latin	Expected Romance	Greek	Modern Greek
SINGULAR	*	*	*	*
nominative	-a	-a	-ê	-ê
vocative	-a	-a	-ê	-ê
accusative	-am	-a	-ên	-ê
genitive	-ae	-e	-ês	-ês
dative	-ae	-e	-êi	*
ablative	-a:	-a	*	*
PLURAL	*	*	*	*
nominative	-ae	-e	-ai	-es
vocative	-ae	-e	-ai	-es
accusative	-as	-as	-as	-es
genitive	-arum	-aru	-ôn	-ôn
dative	-is	-is	*	*
ablative	-is	-is	*	*

As can be seen, Greek preserved its *ê*-stem declension with few changes; the loss of final *-n* in the accusative singular is due to regular sound changes. The situation in the plural is slightly less clear-cut, with a new ending *-es* (originally marking the nominative and vocative plural of Consonant-stem nouns) taking the place of the old *-ai* and *-as* endings.

In contrast, in Romance, while the opposition in the feminine singular between *-a* and genitive/dative *-e* is found to this day in Rumanian and a few other Romance varieties, only a single of the four plural endings, *-as*, is attested.¹² Most Romance languages (French, Provençal, Italian, Spanish, Portuguese) use only a single ending, *-a*, in the feminine singular. Thus, in effect, they do not morphologically distinguish case for nouns belonging to this class, only number: (French *la chèvre/les chèvres*, Spanish *la cabra/las cabras*, Italian *la capra/le capre*, ‘the goat/ the goats).

As Green (1991) points out, there does not appear to be any reason why ‘of the doors’ is expressed in early Spanish by *de las puertas* rather than **laro portaro*, the expected outcome of *illarum portarum* in Latin.¹³ To be sure, he continues, morphophonemic alternations would have been created in the language (due to the fact that Latin *porta* was stressed on the first syllable and *portarum* on the second, the expected Spanish outcome would be *puerta* and **portaro*, respectively), but since these would merely replicate in the nominal system alternations of identical origin found in the verbal system (i.e. *muerdo/*

mordemos “I/we bite”), he concludes that this does not explain the non-survival of these forms. *Mutatis mutandis*, the same comment could be made for most if not all Romance languages, (although it must be said that few if any scholars have in fact done so).

Our next comparison is between the evolutions of Latin and Greek masculine nouns:

Table 2. Masculine Latin *u*- and Greek *o*-stems

CASE	Latin	Expected Romance	Greek	Modern Greek
SINGULAR	*	*	*	*
nominative	-us	-us	-os	-os
vocative	-e	-e	-e	-e
accusative	-um	-u	-on	-o
genitive	-i	-i	-ou	-ou
dative	-o	-o	-ôi	*
ablative	-o	-o	*	*
PLURAL	*	*	*	*
nominative	-i	-i	-oi	-oi
vocative	-i	-i	-oi	-oi
accusative	-os	-os	-ous	-ous
genitive	-orum	-oru	-ôn	-ôn
dative	-is	-is	-ois	*
ablative	-is	-is	*	*

The contrast between the evolution of Latin/Romance and Greek is even more blatant when examining Table 2. As can be seen, other than the loss of the dative and the (phonologically regular) loss of final *-n* in the accusative singular, the declension of modern Greek is identical to that of the Classical language. Normal phonological evolution would not have caused any of the Latin endings to merge with one another, and so one might have expected this particular declension class to have survived somewhat better than the class of feminine *a*-stems (compare the “Expected Romance” case-forms of the singular with those of Table 1).

This in fact is not what we find: only in Old French, Old Provençal and some neighboring varieties do we find the distinction between two cases — nominative and accusative — to have survived in the case of this particular declension class. The Old French forms of *murs* ‘wall’ (from Raynaud De Lage 1966) are given below, with the forms of the Old French definite article

(The Old Provençal forms are highly similar):

	singular	plural
nominative	<i>li murs</i>	<i>li mur</i>
accusative	<i>le mur</i>	<i>les murs</i>

The situation is even simpler in the case of the various other Romance varieties, such as Italian, Spanish, Portuguese, Sardinian or Rumanian¹⁴: in all of these we find that only the accusative singular and the accusative plural have survived (i.e. Italian *il muro/i muri*, Spanish *el muro/los muros*, ‘the wall/the walls’) In short, most of the case forms of Latin, whose distinctive forms would not have been eroded through the effects of regular sound changes, did not survive in any Romance language. With the exceptions noted above, Romance languages did not distinguish any case morphologically for nouns of this class, only distinguishing number, making use of the former accusative.

Table 3. Consonant stems in Latin and Greek

CASE	Latin	Expected Romance	Greek	Modern Greek
SINGULAR	*	*	*	*
nominative	rex	res	phylax	phylakas
vocative	rex	res	phylax	phylaka
accusative	regem	rege	phylaka	phylaka
genitive	regis	regis	phylakos	phylaka
dative	regi	regi	phylaki	*
ablative	rege	rege	*	*
PLURAL	*	*	*	*
nominative	reges	reges	phylakes	phylakes
vocative	reges	reges	phylakes	phylakes
accusative	reges	reges	phylakas	phylakes
genitive	regum	regu	phylakôn	phylakôn
dative	regibus	regibus	phylaksi	*
ablative	regibus	regibus	*	*

Table 3 differs somewhat from its predecessors in presenting the declension of an actual noun (‘king’ in Latin; ‘watchman’ in Greek) rather than mere case endings. This is due to the fact that it is somewhat more difficult to separate the suffixes from the stems. It can be seen that, except for the merger of the accusative and the ablative in the singular, all the case forms which were distinct in Latin would have remained so in Romance. The Greek forms have,

on the surface, changed more than in the case of other declension classes, but in fact, if we discount the remodeling of the nominative/vocative singular forms, the modern Greek forms differ from the Classical solely in (1) substituting the accusative for the genitive in the singular, and (2) substituting the nominative/vocative for the accusative in the plural (perhaps motivated by a desire to avoid confusion with the new nominative singular).

And what is the situation in Romance? Virtually the same as that for *u*-stem nouns: Old French and Old Provençal keep nothing more than the nominative and the accusative in the singular. More surprisingly, they also recreated a nominative/accusative distinction in the plural (on the model of *u*-stem nouns); thus, quoting Old French forms again:

	singular	plural
nominative	<i>li reis</i>	<i>li rei</i>
accusative	<i>le rei</i>	<i>les reis</i>

All other Romance languages, again, fail to distinguish case morphologically, and keep only the former accusative singular and accusative plural to mark number (Italian *il rege/i regi*, Spanish *el rey/los reyes* ‘the king/the kings’).

There existed two other declension classes in Latin, but since in the course of the changes which transformed Latin into early Romance all of its members were assimilated to one of the three classes listed above, they may safely be omitted from this discussion.

What this comparison reveals is that Latin/Romance is extraordinarily analytical when compared to Greek: it has eliminated most of its nominal case-marking morphology in a millennium, whereas Greek, in two millennia, has kept most of its case-marking morphology intact. The contrast between the two languages, from this point of view, is most remarkable.

It should be added that, when one broadens the picture by bringing in data from other Indo-European languages, it is clear that Latin/Romance, not Greek, is the odd man out. Thus, as Martinet (1955: 210) points out concerning Old Irish (a language almost exactly contemporaneous with the earliest attested forms of Romance, i.e. with languages which had either lost case-marking morphology altogether or had reduced the Latin system to a two-case system):

Il est intéressant de noter que le *v[ieil].-irl[andais]. fer* ‘homme’, équivalent sémantique et étymologique du lat. *uir*, distingue dans sa flexion entre sept formes phonologiquement distinctes, tout comme *uir* lui-même.

(It is interesting to note that Old Irish *fer* ‘man’, the semantic and etymological equivalent of Latin *uir*, has seven phonologically distinct case-forms, as does *uir* itself. [trans. by the author])

Slavic, a language family for which we have written documents stretching from roughly a thousand years ago to the present (the same length of time separating Latin from the first texts in Romance languages), has preserved its case-marking morphology largely intact in most of its daughter languages to this very day (Comrie & Corbett 1993).

In the case of Germanic we find a language family whose system of case-marking morphology collapsed long ago. This, however, is easily explained by taking into account the fact that final vowels were typically merged or dropped in Germanic due to strong word-initial stress, and that most earlier case-endings were vocalic (Goyette 1997). Thus, if one examines the declension of Old English *dæg* ‘day’ — singular: nominative/accusative *dæg*, genitive *dæg-es*, dative *dæg-e*; plural nominative/accusative *dæg-as*, genitive *dæg-a*, dative *dæg-um*) (forms from Lass 1994) — and bears in mind that all Old English unstressed vowels and final nasal consonants were dropped in the transition to Modern English, the outcome — singular *day*, genitive singular *day’s*, plural *days*, genitive plural *days’* — is very much what one would expect (leaving aside the analogical extension of original nominative/accusative plural *-s* marker to all plurals, and of the original genitive singular *-s* marker to the plural). Importantly, where this loss of final vowels did not occur in Germanic, as for example in Icelandic (Thráinsson 1994), case-marking morphology remains present.

Returning to Greek, there is one aspect of the modern language which is indubitably a conservative trait and which, for this reason, was not taken into account in the above tables. Both Latin and Classical Greek had three grammatical genders: masculine, feminine and neuter. Whereas in Latin/Romance the neuter gender was eliminated, mostly absorbed by the masculine (leaving various traces in some Romance languages), in Greek all three genders are still intact. At first glance one might be inclined to ascribe this Romance morphological simplification to creolization; however, considering how frequent the reduction of three-gender systems to two-gender (masculine-feminine) systems is in Indo-European (*inter alia*, it is found in Baltic, Celtic, and Central Indo-Aryan [Hindi-Urdu, Punjabi, Romani, et al.]), its presence in Romance need not be explained as being due to anything except normal language change. Conversely, the preservation, as in Greek, of all three genders, is

much less frequent: it is the norm in only one branch of Indo-European, Slavic, and is also found in German and Marathi (mentioned earlier as a conservative Indo-Aryan language as regards to morphosyntax.)

3.2. *Further evidence for the creolization of Proto-Romance*

Before ending this section, however, the reader might ask why the comparison was made with the attested forms of various Romance languages, rather than with a reconstructed Proto-Romance. For one, this was because in the field of nominal morphology in particular, there is at present relatively little scholarly consensus on the form of Proto-Romance.

Furthermore, however, a recent argument strongly suggests that Proto-Romance was in fact quite similar to modern Romance in its paucity of morphology. While most romanists, on the basis of the Rumanian and French data, would reconstruct two or perhaps three case-endings for Proto-Romance, De Dardel & Wüest (1993) propose to explain the data in a rather novel fashion.

They observe that while most Romance languages are similar in retaining forms of the Latin accusative singular and plural, case endings in Rumanian and the Gallo-Romance languages (Old French and Old Provençal) differ quite sharply from one another, and furthermore contain (especially in pronouns) a very large number of analogical forms also revealing profound divergences between Gallo-Romance and Rumanian. This is in sharp contrast to the situation in “caseless” Romance languages, whose nominal forms, including pronouns, are very similar to one another. De Dardel & Wüest propose that the “caseless” Romance languages in fact represent the earliest Romance type, born at the very beginning of Roman expansion, and that the case-marking morphology of Gallo-Romance and Rumanian, far from being directly inherited from Latin, was in fact borrowed from Latin, separately in both cases. According to this thesis, morphological case in Rumanian and Gallo-Romance, far from being an archaic Latin survival, is in fact an indication of contact-induced language change similar to decreolization.

This, in their opinion, would explain the aforementioned facts: for if Proto-Romance was in fact originally a case-marking language, it would be odd for the languages which have in fact preserved case-marking (Rumanian and the Gallo-Romance languages) to differ from one another more than those languages which have abandoned case-marking morphology altogether (al-

ways, curiously, keeping the accusative forms, never the nominative ones). It would also, in their opinion, explain why Sardinian, in most respects the most archaic Romance language, as we saw above, is one of the “caseless” Romance languages.

A most powerful argument they bring to bear in support of their thesis is the fact that the set of prepositions used to mark basic case relations is basically the same in all Romance languages: *ad* as a marker of the dative, *de* or *ad* as a marker of the genitive,¹⁵ and zero or *ad* as marker of the accusative. Had each individual Romance language lost its case-marking system separately, we would doubtless find a greater variety of prepositions used to replace the lost endings.¹⁶ For example, in Germanic, a language family where we know case distinctions to have been lost separately in each language (cf. the changes separating Old from Modern English, given above), English marks its genitive with *of* and its dative with *to*, whereas Dutch uses the (unrelated) prepositions *van* and *aan*, respectively.

One might, of course, argue that the spread of these prepositions throughout Romance-speaking Europe was due to diffusion, perhaps caused by the population movements mentioned by Bonjour, Offler & Potter (1952) cited above as a consequence of the fall of the Roman Empire. However, two objections must be raised: 1) There is evidence that such population movement was as a rule local in nature, and indeed that there was a far greater degree of material and cultural continuity during this period of time than had been hitherto realized (Banniard 1989); and 2) There is no evidence whatsoever for any extensive dialect mixture throughout the Romance language area, as we would expect to be the case had diffusion been the cause of the universal use of *ad* and *de*. In particular, the sound-changes which separate the Romance languages from one another are quite clear-cut, in sharp contrast to a family such as Indo-Aryan, for example, where, as Masica (1991: 458) puts it, regarding the historical phonology of the various languages,

Whatever data does not fit a particular rule must be attributed to borrowing or substratum or otherwise explained. The exceptions are typically so numerous that it is often hard to tell which cases constitute a rule and which are the exceptions.

The most powerful evidence in favor of De Dardel & Wüest’s thesis, however, is a point they themselves fail to emphasize and whose significance seemingly eludes them: throughout Romance, whenever case is lost, the sole remaining form is that of the accusative. This is a highly anomalous situation inasmuch

as, when case is lost in other Indo-European languages, the sole remaining form is normally the nominative, not the accusative. Thus, in Welsh, a language which lost all its case-marking morphology prehistorically, it is quite clear, on the basis of the rules of mutation,¹⁷ that the modern forms of the noun derive from the former nominative, and not from the accusative. For example, the *t/d* mutation found in Modern Welsh *dyn teg* “fair man” and *merch deg* “fair girl” (adjectives follow the noun in Welsh: the mutation is conditioned by the grammatical gender of the head noun in each of the phrases: *dyn* is masculine and *merch* feminine) is explicable through the fact that the original forms were **donjos tekos* and **merka teka*, and that intervocalic consonants were voiced irrespective of word boundaries. Thus, **merka teka* became **merka dega* but **donjos tekos* became **donjos tegos*. Subsequently, with the fall of final syllables, the *t/d* alternation became morphophonemic rather than allophonic. However, it could not have arisen if Welsh nouns were formally derived from the accusative: both masculine and feminine nouns would have ended in the same segment (**-m* or **-n*), meaning that the initial segment of the following adjective would have remained identical, regardless of the gender of the noun.

In Bulgarian, which along with its close relative Macedonian is the only Slavic language to have lost its case-marking morphology, we find that the sole form of nouns remaining in the language today must derive from the nominative rather than the accusative: Bulgarian feminine nouns, such as *žena* ‘woman’, must derive from the Old Church Slavonic nominative: the accusative form would have yielded **ženă* in modern Bulgarian (Carlton 1990).

Thus, the pan-Romance retention of the accusative (singular and plural) as the sole form of the noun goes quite counter to what is found elsewhere in Indo-European. It defies belief that such an unusual change could have taken place independently more than once throughout the Roman Empire, and if we reject the possibility of diffusion (see above), we are left with only one possibility: that the loss of distinct case forms for nouns, and the extension of the former accusatives as the sole noun forms, only took place once, in the beginnings of the expansion of the Roman Empire, as De Dardel & Wüest hypothesize.

4. Verbal Morphology

An examination of Latin and Greek verbal morphology is considerably more difficult than one of nominal morphology. Not only is the number of categories of verbal morphology (finite and non-finite forms; tense, mood, and voice; person and number) considerably greater, but there are also far fewer similarities between the two languages. Indeed, as one authority puts it, “the G[reek] and L[atin] verb systems could not have evolved more differently from their common starting-point if that had been some conscious purpose” (Sihler 1995: 444).

There is no simple one-to-one correspondence between the various Latin and Greek tenses and moods, and indeed, even forms which bear the same label in the two languages often differ as to their actual meaning. To quote but one example: both Latin and Greek have a verbal tense in the indicative called the “pluperfect”. However, whereas in Latin this tense indicates, as in English, an event anterior to another event in the past (i.e. *cado* “I fall/am falling” [present], *cecidi* “I fell” [perfect], *cecideram* “I had fallen” [pluperfect]), in Greek (examples from Ragon 1961) this tense indicates a past action, the consequences of which are not with us as we speak (i.e. *piptô* “I fall/am falling” [present], *peptôka* “I fell” [and still am on the ground] [perfect], *epeptôkein* “I’d fallen” [and still was on the ground] [pluperfect]). It will be noticed, incidentally, that the term “perfect” also indicates different tenses in the two languages. For this reason the following comparison, contrary to the one in the preceding section, will confine itself, though with regret, to one particular phenomenon: the fate of the synthetic passive¹⁸ in both languages, on the one hand because of the far-reaching consequences of this morphological change, on the other because the passive forms in Latin and Greek were, semantically speaking, similar. A thorough comparison of the verbal morphologies of both languages would require far more space than is available here.

Two actual verbs, both regular, will be used, Latin *amare* ‘to love’ and Greek *lyein* ‘to loosen’ to exemplify the morphological alternation discussed:

Table 4. Indicative present active/passive morphology in Latin and Greek

PERSON/ NUMBER	Latin indicative present active	Latin indicative present passive	Greek indicative present active	Greek indicative present passive
1st singular	amo	amor	lyô	lyomai
2nd singular	amas	amaris	lyeis	lyei
3rd singular	amat	amatur	lyei	lyetai
1st plural	amamus	amamur	lyomen	lyometha
2nd plural	amatis	amamini	lyete	lyesthe
3rd plural	amant	amantur	lyousi	lyontai

(The forms of the dual in Greek have not been included, for the same reason as in the nominal morphology section.)

As can be seen, both languages resorted to a distinctive set of endings to express the passive voice. It should be pointed out that not only were the Greek endings more ambiguous (inasmuch as the active third singular and passive second singular forms were identical), but their degree of variation (from one tense/mood to the next) was greater, as can be seen in Table 5:

Table 5. Indicative imperfect active/passive morphology in Latin and Greek

PERSON/ NUMBER	Latin indicative imperfect active	Latin indicative imperfect passive	Greek indicative imperfect active	Greek indicative imperfect passive
1st singular	amabam	amabar	elyon	elyomên
2nd singular	amabas	amabaris	elyes	elyou
3rd singular	amabat	amabatur	elye	elyeto
1st plural	amabamus	amabamur	elyomen	elyometha
2nd plural	amabatis	amabamini	elyete	elyesthe
3rd plural	amabant	amabantur	elyon	elyonto

Comparing these verbal forms to those in Table 4, it will be noticed that, formally speaking, the forms of the passive in the indicative imperfect are much closer to the forms of the indicative present in Latin than in Greek: whereas in Latin one could derive the passive from the active by a single set of rules, valid for all persons of the indicative present and imperfect save the first singular,¹⁹ this would be quite impossible in Greek, as a simple glance at the forms in Tables 4 and 5 shows. Thus, one must agree with Green's statement

that “...when viewed purely as morphology, the [Latin] passive paradigms are remarkably solid and efficient” (1991: 85).

This being the case, it is all the more remarkable that passive morphology, in the transition from Latin to Romance, disappeared without leaving a trace (Elcock 1960). This change is remarkable: for as Green again points out, these forms were for the most part in no danger of being eliminated through the operation of sound changes. The consequences of this loss were profound, for since most finite verb forms in the active had a corresponding inflected passive form, what happened, in effect, was that a huge piece of Latin finite verb morphology was eliminated. In its place Romance languages use various analytical forms.

In contrast, Modern Greek still preserves its passive-marking morphology (albeit in a considerably remodeled form). Nor is it unusual or archaic among Indo-European languages in so doing. To be sure, there is no evidence for any synthetic passive morphology in Slavic (Schenker 1993), and in Germanic, an inherited synthetic passive is attested only in Gothic, its stages of loss, if any, unknown (these forms coexisted with analytical ones similar to those elsewhere in Germanic; whether they would have survived had Gothic persisted as a living language is an open question [Lehmann 1994]). More to the point, however, Old Irish passive morphology was alive and well in that language even while it was long gone in Romance (Lehmann & Lehmann 1975), and synthetic passive forms directly derived from Indo-European forms are found in Modern Celtic and Albanian (Bubenik & Hewson 1997).

Again, then, as in the case of nominal morphology, we find that Latin/Romance evolved in a considerably more analytical direction over a single millennium than Greek did over two. Furthermore, we again see that this phenomenon is due to the innovative nature of Latin/Romance, not the conservative nature of Greek, as a brief examination of other branches of Indo-European shows.

5. *Varia*

As the two preceding sections have shown, Latin/Romance has eliminated, in the course of its thousand-year history, most of its nominal case-marking morphology and a considerable portion of its verbal morphology. What is of interest for our purposes is that the lost forms were replaced in function by

various analytical means of expression (prepositions and word order to mark case in nouns; periphrastic passive constructions using the verb ‘to be’ + past participle and various other devices to replace passive morphology on verbs), but never by newly-created synthetic morphological markers. The contrast with Greek, incidentally, is all the more remarkable when one bears in mind that during the relevant period the two languages were in close contact.

Of far greater interest, however, is the remarkable fact that during this period, Late Latin/Early Romance did not grammaticize any of its many analytical structures: this fact suggests a type of linguistic transformation more akin to creolization than to normal linguistic change.

This statement may come as something of a surprise to a number of readers, since two of the best-known cases of grammaticization — the “future/conditional” synthetic forms and the adverbial suffix *-mente* — are often said to be “Romance”. However, it should be clarified that this is not quite true *strictu sensu*.

The synthetic future and conditional were indeed originally analytical formations, consisting of the infinitive followed by the verb ‘to have’ in the present or past tense, respectively: French (*je*) *chanterai* ‘I will sing’ < *chanter* + *ai*, ‘to sing’ + ‘I have’; likewise (*je*) *chanterais* ‘I would sing’ < *chanter* + (*av*)*ais*, ‘to sing’ + ‘I had’. However, it should first be noted that these developments are by no means pan-Romance, being totally absent from Sardinian, Southern Italian varieties, Rumanian and some varieties of Rhaeto-Romance. Furthermore, their grammaticization — in those languages where they are found — had barely begun in many languages when the first written records appear, and indeed is not quite finished in the case of at least one language, (Continental) Portuguese, where object clitic pronouns may still be inserted between the infinitive and its “ending”, i.e. *dir-se-ia* “It would be said” (literally, “it would say itself”); compare Spanish *se diría*, Italian *si direbbe*.

It should also be pointed out that the Italian conditional differs that of its Romance sisters in that its endings go back to a cliticized form of the verb ‘to have’ in the perfect tense, whereas in other Romance languages with a synthetic conditional the endings go back to a cliticized form of the verb ‘to have’ in the imperfect tense; thus, the above forms *diría* and *direbbe* cannot be traced to a common Romance, or indeed Latin, form. This is further proof that the grammaticization of the auxiliary verb is not a pan-Romance innovation. Clearly, the grammaticization of these auxiliaries cannot be taken to have

occurred during the Early Romance period, but rather much later, after the separation of the various Romance languages.

The story of the adverbial suffix *-mente* is quite similar to that of the synthetic future/conditional forms: it is absent from Sardinian, Rumanian and Southern Italy as well; in those languages where it is present its grammaticization is to this day not necessarily quite complete. Thus, ‘clearly and openly’, in Spanish, would be *clara y abiertamente* (‘clear and openly’), with *-mente* turning both adjectives into adverbs (contrast French *clairement et ouvertement*). Again, we must suppose the grammaticization of *mente* to have occurred after the separation of the various Romance languages.

Further evidence that the restructuring of Latin/Romance is more akin to creolization than to normal linguistic change is supplied not only by comparison with Greek and other Indo-European languages, but by the later history of the Romance languages themselves. Leaving aside those changes due to the operation of regular phonological change (such as the disappearance of most person-marking suffixes in French verbs), we nowhere across the Romance languages find anything like the radical restructuring which occurred in its Latin/Romance ancestral stage. In nominal morphology, the elimination of the two-case declension in French and Provençal is not surprising when we bear in mind that two cases and two numbers had to be expressed by only two endings, *-s* and zero. In verbal morphology, the most dramatic change (most conspicuous in French but present in several other languages as well) has been the loss of the perfect and its morphologically related subjunctive imperfect, a formal simplification of verbal morphology to be sure, but nowhere near as dramatic as the total loss of all synthetic passive paradigms, and in any case not even found in other Romance languages. This is in stark contrast to the large-scale loss of morphology which occurred in the transition from Latin to Romance.

Conversely, while there is no documented grammaticization at the Latin/Proto-Romance stage, grammaticization of formerly independent elements has occurred and is an ongoing process in many modern Romance languages. In the verbal system, many auxiliary verbs are on their way to becoming pure person-marked tense or mood markers (Champion 1980; see Steinkrüger 1995 for a detailed examination of this process in Catalan), and pronominal clitics throughout Romance are well on their way to becoming verbal agreement morphemes (Auger 1995, Uritescu 1997). Likewise, regular sound change has not always led to morphological simplification: as a result of its attrition of

final consonants, French has created a distinction between strong and weak masculine adjectival forms (Morin 1992). This, again, contrasts with the fact that virtually no grammaticization appears to have taken place in the transition from Latin to Romance.²⁰

The transition from Latin to Romance, where massive morphological loss and virtually no grammaticization occurred, is also anomalous when compared to the changes which separate Latin from its Indo-European sister languages. Whereas some loss of morphological distinctions did indeed occur in the history of Latin (for example, of a seventh case form, the locative, preserved in Osco-Umbrian [Buck 1974]), there also occurred a good deal of grammaticization: in the verb system, for instance, of the six finite tense forms in the indicative (using the first person singular of the verb *amare* 'to love' to exemplify: present: *amo* 'I love', perfect *amavi* 'I loved', imperfect *amabam* 'I was loving', pluperfect *amaveram* 'I had loved', future *amabo* 'I shall love', future anterior *amavero* 'I will have loved'), only the first two tenses are inherited from Indo-European. The other four are later internal Latin creations, the endings being, historically, cliticized and ultimately grammaticized forms of the copula (originally post-verbal, in line with the verb-final syntax of Early Indo-European; cf. Lehmann 1993).

Thus, whereas loss of synthetic forms was accompanied by grammaticization of analytical forms in both the passage from Indo-European to Latin as well as the passage from Early to Modern Romance, in accordance with the predictions of Schwegler (1990), in the passage from Latin to Early Romance, there was a massive loss of morphological marking without any accompanying grammaticization of analytical forms. This clearly suggests that something other than normal linguistic change took place during this period, and postulating creolization as an explanation is in accordance with both the linguistic and the sociopolitical facts.

In short, it would indeed appear that creolization did indeed play a role in the emergence of Romance from Latin.

6. Conclusion

This paper sought to establish whether there is any linguistic basis supporting the contention that creolization played a significant role in the emergence of Romance from Latin, a possibility which cannot be lightly dismissed in light

of the external history of Latin. Based on the premise that creolization differs from normal language change in leading to a radically heightened degree of analyticity, the analysis utilized Greek as a “yardstick” indicating “normal”, gradual linguistic change, in order to ascertain whether the evolution from Latin to Romance was in fact more “creole-like” than that of Greek. Data from other Indo-European languages was also taken into consideration, in order to guard against the possibility of idiosyncratic changes in the linguistic history of Greek.

It was found that in nominal as well as in a major subset of verbal morphology, Latin evolved in the direction of much more radical analyticity than Greek, and moreover did not, during the relevant period, create any new synthetic structures, in contradistinction to what is expected in the case of “normal”, gradual linguistic evolution.

On this basis, postulating creolization as a contributing factor in the evolution of Latin/Romance becomes a matter of explanation rather than speculation, and indeed would appear to be the simplest (and therefore, on the basis of Occam’s razor, should be the one accepted), as otherwise we are left with having to explain the greater analytical aspect of Romance when compared to Greek in the nominal as well as in the verbal system.

If one accepts the validity of the data presented here, then we must observe before closing that a refutation of the theory presented here, to be taken seriously, would need to present an alternative explanation as to why, in contradistinction both to its Romance daughters and its Indo-European sisters, the transition from Latin to Romance was accompanied by such a radical, unidirectional shift in the direction of analyticity.

Finally, it is to be hoped that the methodology presented here will be applied to other putative cases of “creoloid” languages or language families: the possibility that there may have been other hitherto-undetected cases of creole-influenced evolution should certainly not be dismissed out of hand. It must be remembered, however, that few of the languages of the world are anywhere near as well-studied as Romance, and the overwhelming majority do not have written records going as far back in time: this may very well mean that we will have to be content to say, regarding the role creolization may have played in the history of most of them, *ignoramus et ignorabimus*.

Notes

1. I would like to thank the following people for their input: Armin Schwegler, who commented on the various ideas defended herein over electronic mail with extraordinary patience and rigor, and who encouraged me to present them; Galina Alexandrova, Dawn Harvie and James Walker, who listened to a preliminary version of my presentation and made some valuable suggestions; the audience at the January 1997 meeting of the Society for Pidgin and Creole Linguistics, whose questions and comments indubitably proved helpful; John McWhorter, for his encouraging remarks; an anonymous reviewer for helpful comments; and, last but definitely not least, Konrad Koerner for his unwavering support. Remaining infelicities, of both style and substance, are of course my responsibility alone.
2. The terms “substratum” and “superstratum” are here used in the sense understood by historical linguists rather than creolists: whereas the latter use the two terms to refer to the “parent” languages of a creole, the former use these terms to refer to languages which have influenced the development of another language, so that, for example, when historical linguists refer to a Gaulish “substratum” and a Frankish “superstratum” in French, they do not mean that French was born because of contact between these two languages: rather, they mean that French was influenced by both languages. Had such influence never occurred, there would still exist a Romance language in Northern France. By contrast, when creolists, for example, speak of Haitian Creole as having a French “superstratum” and a Fongbe “substratum”, they mean that Haitian Creole was born because of contact between these two languages, and that had there never been a substrate and a superstrate language, there would be no such thing as a Haitian Creole language.
3. It should be remembered that the Romance-speaking area, in Europe, has shrunk considerably since the fall of the Western Roman Empire: Bavaria, Austria, German-speaking Switzerland, Flanders and most of the Netherlands, Slovenia, Croatia and North Africa were Romance-speaking originally, and the same may be true of other parts of the Western Roman Empire as well.
4. As well as from other early Indo-European languages spoken in what was to become Romance-speaking territory as well, such as Oscan, Umbrian, Gaulish, Ibero-Celtic, and Lepontic, which are considerably closer, typologically, to Latin than to Romance (see Buck [1974] for the first two and Ball [1992] for the latter three). This clearly indicates that one cannot explain the birth of Romance from Latin as a case of substrate influence (see footnote 2).
5. The term “Vulgar Latin” shall be avoided in this paper, as its exact meaning varies greatly from author to author, few of whom (Elcock 1960 is a rare exception) precisely define what they mean by this label.
6. Similar such criticism can be levelled at the other works cited; this is particularly true of Southworth (1971), whose theory that Marathi, contrary to other Indo-Aryan languages, was originally a creolized language is extraordinarily difficult to support in light of the fact that Marathi is in many respects a remarkably archaic form of Indo-Aryan, particularly with regards to morphosyntax (for details, see Masica 1991).
7. The expansion of Greek as a prestige language throughout the Mediterranean and Near East might seem to belie this claim, but this linguistic “expansion” of Greek was

primarily an elite phenomenon, leaving the bulk of the non-Greek-speaking peoples unaffected; this contrasted sharply was the expansion of Latin/Romance, as Meillet (1928b) points out.

8. See Banfi (1993) for an explanation of the differential rates of language change in Latin and Greek as being due to such differences in sociopolitical history.
9. Please note that the converse is not true: if we were to find the evolution of Italic Greek dialects to be more reminiscent of Romance than Balkan Greek, this could well be explained as being due to the influence of Romance, which has been influencing Italic Greek for well over a millenium and whose prestige today is such that it seems likely that Italic Greek will become extinct in the near future, through shift of its speakers to surrounding varieties of Italian (Profili 1985).
10. The following should be noted regarding the romanization of Greek: the letters η and ω will be transcribed by \hat{e} and \hat{o} respectively; the letter υ by y except if it is the second element of a diphthong, in which case u will be used; the letters ζ , ξ and ψ by dz , ks and ps respectively; θ , ϕ and χ by th , ph and kh respectively; *spiritus asper* by h ; and subscript ι by post-vocalic i ; stress will not be marked. This schema will be used for both Ancient and Modern Greek forms, in spite of the phonetic values of the letters having often changed over time: thus, for example, ϕ will always be rendered as *ph*, in spite of the fact that it was an aspirate consonant in Ancient Greek and is a fricative in the modern language. The reader should also be cautioned that this schema, being a transliteration rather than a transcription, does not indicate some mergers which have taken place in the history of Greek and which are not indicated in the spelling: for example, the letters transcribed here as i , y , and \hat{e} , along with most diphthongs having i as a second member, are all realized identically in modern Greek as $/i/$ (a phenomenon known as “itacism”).
11. It will be noticed that Greek also differed from Latin in that it had a definite article, which Latin did not have, in contradistinction to its Romance descendants.
12. Earlier generations of romanists believed that the plural endings *-i* and *-e* of Italian and Rumanian (i.e. Italian *lupo/lupi* ‘wolf/wolves’, *casa/case* ‘house/houses’) were derived from the Latin nominative plural forms, as opposed to the Spanish or French forms in *-s* (i.e. Spanish *lobo/lobos*, *casa/casas*), which were held to derive from the Latin accusative. Following Seklaoui (1989) and Maiden (1996), all these forms are assumed to be in fact derived from the Latin accusative, with a sound-change whereby final *-s* was vocalized to *-j* (followed by monophthongization of the resulting diphthongs), accounting for the Italian and Rumanian forms.
13. Some Romance scholars will be quick to point out that, as a matter of fact, reflexes of the masculine genitive plural of the demonstrative pronoun, *illorum*, have indeed survived into at least some varieties of Romance, i.e. French *leur*, Catalan *llor*, Italian *loro*, Rumanian *lor*. This fact is not denied here; however, wherever they are found, they have been lexicalized as third-person plural possessive adjectives, (i.e. French *leur maison* ‘their house’) and thus may only be said to continue the Latin genitive plural formally, not semantically. They do not stand in anything like a paradigmatic relationship to what in Latin were their corresponding nominative forms, and thus cannot be said to be, synchronically, genitive plurals.
14. It is possible that Rumanian masculine vocative forms in *-e* continue the Latin vocative, but since a) such forms are also found in Slavic, and b) it is uncontroversial that

- Rumanian borrowed its feminine vocative suffix *-o* from Slavic, it would seem simpler to consider its masculine vocative suffix to also be a Slavic borrowing.
15. No modern Romance standard language uses reflexes of *ad* to mark the genitive, but such usage is found in many non-standard varieties, cf. Colloquial French *le sac à la femme* 'the woman's bag' (versus Standard French *le sac de la femme*).
 16. An example of such an internal development is Rumanian, which formerly marked its accusative nouns with zero but now marks its animate accusatives by means of a preposition *pe* (Latin *per*) (Rosetti 1986).
 17. For an account of the rise of Celtic Mutation and its failure to appear elsewhere in Indo-European, see Goyette (1998).
 18. This term, although traditional, is somewhat misleading, inasmuch as in Latin, for example, intransitive verbs could bear "passive" morphology, i.e. *currit* '(s)he runs/is running', *curritur* '(someone) runs/is running'. Perhaps "impersonal" would be a better label.
 19. Such a set of rules (*-s* becomes *-ris*, *-t* becomes *-tur*, *-mus* becomes *-mur*, etc.) is in fact often presented in pedagogical grammars of Latin.
 20. One thus cannot thus follow Bennett's (1980) attempt to see in the transition from Latin to Romance a type of linguistic change similar to that from Standard French to modern colloquial Parisian French.

References

- Auger, J. 1995. "Les clitiques pronominaux en français parlé informel: une approche morphologique". *Revue Québécoise de Linguistique* 24: 21-60.
- Bailey, C.-J. N. and Maroldt, K. "The French lineage of English". In Meisel: 21-53.
- Ball, M. (ed.) 1993. *The Celtic languages*. London: Routledge.
- Banfi, E. 1993. "Latin System's *enérgeia* vs. Greek system's *argia*: A problem of linguistic change theory". In *Historical Linguistics 1989: Papers from the Ninth International Conference on Historical Linguistics*, H. Aertsen and R.J. Jeffers (eds), 47-58. Amsterdam: John Benjamins.
- Banniard, M. 1989. *Genèse culturelle de l'Europe (Ve -VIIIe siècle)*. Paris: Le Seuil.
- Bennett, W.A. 1980. "Language populaire (sic) and language drift". *Neophilologus* 14:321-32.
- Bonjour, E., Offler, H.S. and Potter, G.R. 1952. *A short history of Switzerland*. London: Clarendon.
- Bubenik, V. and Hewson, J. 1997. *Tense and aspect in Indo-European languages*. Amsterdam: John Benjamins.
- Buck, C. 1974. *A grammar of Oscan and Umbrian*. Hildesheim, New York: Olms.
- Carlton, T.R. 1990. *Introduction to the phonological history of the Slavic languages*. Columbus: Slavica.
- Champion, J. 1980. "A note on prefixed inflection in Romance". In *Italic and Romance: linguistic studies in honor of Ernst Pulgram*, H.J. Izzo (ed.), 97-102. Amsterdam: John Benjamins.

- Chaudenson, R. 1983. "Où l'on reparle de la genèse et des structures des créoles de l'Océan Indien". *Études Créoles* 6, 157-224.
- Chaudenson, R. 1995. *Les créoles*. Paris: Presses Universitaires de France.
- Comrie, B. (ed.) 1987. *The world's major languages*. New York: Oxford University Press.
- Comrie, B. and Corbett, G. (eds) 1993. *The Slavonic languages*. London: Routledge.
- De Dardel, R. and Wüest, J. 1993. "Les systèmes casuels du proto-roman: Les deux cycles de simplification". *Vox Romanica* 52: 25-65.
- Elcock, W. D. 1960. *The Romance languages*. London: Faber & Faber.
- Garnsey, P. and Saller, R. 1987. *The Roman empire: Economy, society and culture*. Berkeley and Los Angeles: University of California Press.
- Goyette, S. 1997. "Why English ain't no creole". Paper presented to the Society for Pidgin and Creole Linguistics, London, England.
- Goyette, S. 1998. "The origins of Celtic mutation". *Buffalo Papers in Linguistics*, ed. by Cori Grimm, Eve Ng and Julie Olen, 47-60. State University of New York, Buffalo Department of Linguistics.
- Green, J. N. 1991. "The collapse and replacement of verbal inflection in Late Latin/Early Romance: how would one know?" In Wright, 83-99.
- Holm, J. 1988. *Pidgins and creoles (Volume 1)*. Cambridge: Cambridge University Press.
- Joseph, B. 1987. "Greek". In Comrie, 144-173.
- König, E. and Van der Auwera, J. (eds) 1994. *The Germanic languages*. London: Routledge.
- Lass, R. 1994. *Old English: A historical linguistic companion*. Cambridge: Cambridge University Press.
- Lehmann, W. P. 1993. *Theoretical bases of Indo-European Linguistics*. New York: Routledge.
- Lehmann, W.P. 1994. "Gothic". In König and Van der Auwera, 19-37
- Lehmann, W. P. and Lehmann, R.P.M. 1975. *An Introduction to Old Irish*. New York: The Modern Language Association of America.
- Leonard, C.S., Jr. 1978. *Umlaut in Romance*. Grossen-Linden, Germany: Hoffmann.
- Maiden, M. 1996. "On the Romance Inflectional endings -i and -e". *Romance Philology* 1: 147-182.
- Masica, C. P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- McWhorter, J. 1998. "Identifying the creole prototype: Vindicating a typological class". *Language* 74: 788-818.
- Martinet, A. 1955. *Économie des changements phonétiques*. Bern: A. Francke.
- Meillet, A. 1928a. *Esquisse d'une histoire de la langue latine*. Paris: Payot
- Meillet, A. 1928b. *Les langues dans l'Europe nouvelle*. Paris: Payot.
- Meisel, J. (ed.) 1977. *Langues en contact - pidgins - creoles - languages in contact*. Tübingen: TBL.
- Monteilhet, H. 1984. *Neropolis: Roman des temps néroniens*. Paris: Éditions du Juillard.
- Morin, Y.-C. 1992. "Un cas méconnu de la déclinaison de l'adjectif français: les formes de liaison de l'adjectif antéposé". In *Le mot, les mots, les bons mots / word, words, witty words: Hommage à Igor A. Mel'čuk à l'occasion de son soixantième anniversaire*, A. Clas (ed.), 233-250. Montréal: Les presses de l'Université de Montréal.
- Mufwene, S. 1988. "Column". *Journal of Pidgin and Creole Languages* 3:265-76.
- Profili, O. 1985. "La romanisation d'un parler grec de l'Italie du Sud par les parlers romans

- environnants". Actes du XVIIème Congrès International de Linguistique et Philologie Romanes (Volume 7), 129-39. Publications de l'Université de Provence.
- Ragon, E. 1961. *Grammaire grecque*. Paris: Éditions De Gigord.
- Raynaud De Lage, G. 1966. *Introduction à l'ancien français*. Paris: Société d'Édition d'Enseignement Supérieur.
- Rosetti, A. 1986. *Istoria limbii române*. Bucarest: Editura Științifică și Enciclopedică.
- Rohlf, G. 1949. *Historische Grammatik der unteritalienischen Gräzität*. Munich: Verlag der Bayerischen Akademie der Wissenschaften.
- Schenker, A. 1993. "Proto-Slavonic". In Comrie and Corbett, 60-121.
- Schlieben-Lange, B. 1977. "L'origine des langues romanes: un cas de créolisation?" In Meisel, 81-101.
- Schwegler, A. 1990. *Analyticity and syntheticity: A diachronic perspective with special reference to Romance languages*. Berlin: Mouton de Gruyter.
- Seklaoui, D. 1989. *Change and compensation: Parallel weakening of /s/ in Italian, French and Spanish*. New York: Peter Lang.
- Sihler, A. 1995. *New comparative grammar of Greek and Latin*. New York: Oxford University Press.
- Southworth, F.C. 1971. "Detecting prior creolization: an analysis of the historical origins of Marathi". In *Pidginization and creolization of languages*, Dell Hymes (ed.), 255-76. Cambridge: Cambridge University Press.
- Steinkrüger, P.O. 1995. "Grammatikalisierungen von Auxiliaren und Copulae im Katalanischen der Decadència". *Zeitschrift für Katalanistik* 8: 35-62.
- Thomason, S. and Kaufman, T. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thráinsson, H. 1994. "Icelandic" In König and Van der Auwera, 142-189.
- Todd, E. 1994. *Le destin des immigrés: Assimilation et ségrégation dans les démocraties occidentales*. Paris: Éditions du Seuil.
- Uritescu, D. 1997. "L'effacement du /l/ en québécois et le type morphologique du français". *Revue de Linguistique Romane* 61: 397-437.
- Versteegh, K. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam: John Benjamins.
- Versteegh, K. 1986. "Latinitas, hellenismos, 'Arabiyya'". *Historiographia Linguistica* 13: 425-448.
- Wright, R. (ed.) 1991. *Latin and the Romance languages in the Early Middle Ages*. London: Routledge.

The Creole Verb

A Comparative Study of Stativity and Time Reference

John Holm et al.*

1. Introduction

This paper represents the first attempt to test the claim that the unmarked creole verb “...signifies past with nonstatives and nonpast with statives” (Bickerton 1979: 309) by examining a wide range of creole languages, including not only Atlantic varieties (Sranan Creole English, Jamaican Creole English, Guyanese Creole English, Krio Creole English, Haitian Creole French, Dominican Creole French, Cape Verdean Creole Portuguese, Guiné-Bissau Creole Portuguese, Angolar Creole Portuguese, Negerhollands Creole Dutch, Papiamentu Creole Spanish, and Palenquero Creole Spanish), but also five non-Atlantic creoles (Tok Pisin Pidgin/Creole English, Nagamese [or creolized Assamese], Nubi Creole Arabic, Seychellois Creole French, and Zamboangueño [a dialect of Philippine Creole Spanish]).

This study is part of a research project on comparative creole syntax. Contributors, who are at the City University of New York unless otherwise indicated, include Lilian Adamson (U. Amsterdam) [Sranan], Dwijen Bhattacharjya [Nagamese], Daniel Chapuis [Dominican, Seychellois], Michel DeGraff (Massachusetts Institute of Technology) [Haitian], Christa de Kleine [Negerhollands], Nicholas Faraclas (U. Papua New Guinea) [Tok Pisin], Kate Green [Palenquero], Gerardo Lorenzino [Angolar], Heliana Mello [Cape Verdean], Abigail Michel [Papiamentu], Jonathan Owens (U. Bayreuth) and Cornelia Khamis (U. Hamburg) [Nubi], Peter Patrick (U. of Essex) [Jamaican], Salvatore Santoro [Zamboangueño], Ronald Simon [Guyanese], Miki Suzuki [Guiné-Bissau] and Sorie Yillah [Krio]. Most are native speakers of

the creole they are describing and/or its superstrate. All examples from a given creole were by the corresponding contributor, via native intuition, field work, or reference to other sources as marked.

We have tried our best to obtain data that are truly comparable, based on what we hope is a common understanding of crucial notions involving tense and aspect (as discussed in Comrie 1976, for example), but we are aware of the complexities and pitfalls of comparative syntax (and the semantics corresponding to that syntax) and realize the potential for limitations in our work, e.g. lack of strictly parallel syntactic or discourse contexts. We propose, however, that it is more constructive to present a preliminary work that will gain new ground for the field rather than renounce such an effort entirely; our work stands for future researchers to build upon.

Our data base includes not only unmarked verbs but also those marked for anterior or past tense. Data are presented according to lexically-based sub-groupings of creoles, first within the Atlantic group and then within the non-Atlantic group. Among unmarked verbs (section 2.0), we examine stative verbs with non-past reference (2.1) and stative verbs with past reference (2.2), as well as non-stative verbs with past reference (2.3) and non-stative verbs with non-past reference (2.4). Among verbs marked for anterior or past tense (3.0), we examined both stative verbs with past reference (3.1) and non-stative verbs with (past-before-) past reference (3.2). Our findings are summarized in section 4; see a note following this for an explanation of the abbreviations used throughout the text.

2. Unmarked Verbs

The notion of ‘unmarked’ verbs in creole languages refers to the absence of a TMA marker (indicating tense, mood or aspect) before verbs, which can be either stative (i.e. indicating a state) or non-stative (e.g. indicating an action). The semantic contrast between marking and non-marking on the one hand, and between stative and non-stative verbs on the other, results in different tense and aspectual meanings, as the following sections show.

Yillah (Holm forthcoming) notes that in Krio CE “The unmarked verb can refer to either the present or the past tense, regardless of whether it describes a state or an action; temporal reference is usually clear from the context.” This appears to be true of Atlantic creoles generally.

2.1 *Statives with non-past reference*

Stative verbs (e.g. ‘love,’ ‘know’) express feelings or states, rather than actions. Stative verbs unmarked for tense usually refer to the non-past in Atlantic creoles, their time reference normally being indicated by the context:

- (1) A *umapikin* ***gersi*** *en* *papa*. (SR)
 DET daughter resemble 3s father
 ‘The daughter resembles her father.’
- (2) I ***lek*** *di* *os* *wo* *Olu mek* *for* *am*. (KR)
 3s like the house REL Olu make for 3s
 ‘She likes the house which Olu built for her.’

Jamaican is typical of Caribbean varieties of Creole English which coexist with standard English and form a continuum of lects extending from a standard-like acrolect to a creole-like basilect. Jamaican verbs which refer to past time are only infrequently inflected with the *-ed* suffix which marks the English past tense, and this is commonly ascribed to “interference” from standard English. The Atlantic creole basilectal pattern of marking past or anterior tense with a preverbal element, such as JC *ben*, is not always present either, however, and past-reference verbs are commonly not marked at all.

Unmarked JC verbs, then, are not part of a privative opposition (cf. Sankoff 1990): they have no unique interpretation as either past or present, anterior or non-anterior. Where an unmarked verb falls within the scope of a contextual element with temporal reference (e.g. a time adverbial, or the clause’s position in a string of sequenced narrative clauses), it acquires that reference.

Patrick (Holm forthcoming) notes that although unmarked stative verbs are often said to be characteristically (Agheyisi 1971) or categorically (Bickerton 1975) interpreted as non-past in reference in Atlantic pidgins and creoles, this is not always the case in Jamaican. Although (3) below refers to the present, (23) in section 2.2 clearly refers to the past.

- (3) *Kieti* ***waan*** *wan neda buk*. (JC)
 ‘Katie wants another book.’ (Bailey 1966: 38)

The pattern in Guyanese CE is similar to JC:

- (4) *Mi na no wai dem a du dis ting.* (GC)
 1S NEG know why 3P PROG do this thing
 'I don't know why they are doing this.' (Bickerton 1975: 29)

In Negerhollands, the tense reference of an unmarked stative verb depends on the context and can be either non-past (as below) or past, as in (25) in section 2.2.

- (5) *Hunder weet sie nest.* (NH)
 chicken knows its nest
 'He knows himself.' (a saying) (Hesseling 1905: 134)

In both Haitian and Dominican Lesser Antillean CF, unmarked stative verbs generally refer to a situation which is concomitant with the time of utterance:

- (6) *Bouki renmen chat la.* (HA)
 Bouki love cat DET
 'Bouki loves the cat.'
- (7) *Kòkòti enmen liv-sala.* (DM)
 Kokoti love book-DEM
 'Kokoti loves this book.'

According to Green (Holm forthcoming), Palenquero CS is unlike other Atlantic creoles regarding the occurrence and distribution of unmarked verbs. In PL the unmarked form is an allomorph of the completive aspect marker *a*. There are predictable environments in which this marker can be deleted. In sentence (8) below, for example, the negator *nu* triggers its deletion:

- (8) *Pogké? pogk' í ø polé pasá jambre nu.* (PL)
 why because I TMA can pass hunger NEG
 'Why? Because I never get hungry.'
 (Friedemann & Patiño Rosselli 1983: 121)

In Papiamentu CS, some (but not all) stative verbs can be unmarked for tense or aspect, according to Michel (Holm forthcoming). PP verbs can be divided into three groups: (a) verbs requiring the non-past marker *ta* (non-statives); (b) verbs requiring either *ta* or \emptyset (some statives); and (c) verbs requiring \emptyset (other statives). Some verbs in the last group never take either *ta* (non-past tense) or *a* (past tense), e.g. *ta* 'to be', *tin* 'to have', *por* 'can, may',

ke ‘to want’, *mester* ‘must, should’:

- (9) *Mi* \emptyset ***ke*** *drumi*. (PP)
 I TMA want sleep
 ‘I want to sleep.’

Certain verbs in (b), however, are sometimes marked for non-past tense, but this indicates that they are functioning as non-stative verbs with meanings that may differ from the stative verbs of the same form:

- (10) *Mi* \emptyset ***yama*** *Pòppy*. (PP) (STATIVE)
 I TMA call Poppy
 ‘I am called Poppy.’
- (11) *Mi* ***ta*** ***yama*** *mañan*. (PP) (NON-STATIVE)
 I NON-PAST call tomorrow
 ‘I will call tomorrow.’

Group (b) also includes some stative verbs that can either be marked or unmarked for non-past reference without any change of meaning, e.g. *debe* ‘to owe’, *gusta* ‘to like’, *kosta* ‘to cost’, *stima* ‘to love’; *merese* ‘to deserve’, etc.

In Cape Verdean CP, stative verbs take a non-past reading when unmarked:

- (12) *Bu* ***quere*** *ba toma lumi*.
 you want go take fire
 ‘You want to go get fire.’ (Parsons 1923 cited by Silva 1985: 145)

This is also the case in Guiné-Bissau CP:

- (13) *Nya irmon* ***misti*** *bay utru tera*. (GB)
 1s brother want go other country
 ‘My brother wants to go to another country.’ (Kihm 1994: 26)

In Angolar CP, stative verbs are always unmarked, taking none of the preverbal markers used with non-statives. The tense and aspect of stative verbs are inferred from the context, as in the following sentence; here it is clear from the discourse that *te* ‘to have’ refers to a permanent state of affairs, i.e. non-past:

- (14) *Angu dhe tha ua txiba ma nɔ* ***te*** *letu tia*. (AN)
 angu this be a banana that we have in land
 ‘The *angu* is a kind of banana that we have in this country.’

In Seychellois CF, stative verbs used without TMA markers generally refer to the time of utterance or to that of the immediate context:

- (15) *Mô ana ê zoli masin.* (SC)
 1s have DET beautiful car
 'I have a beautiful car.' (Corne 1977: 102)

In Tok Pisin, stative verbs are assumed to be non-past, unless marked otherwise by an auxiliary, an adverbial, or by context:

- (16) *Mi save long yu.* (TP)
 I know LOC you
 'I know you.'

In Zamboangueno CS, stative verbs may or may not take the preverbal marker *ta*, which can indicate either present tense or durative (progressive or habitual) aspect. In sentence (17) below *pwede* 'can' is preceded by *ta*, while in (18) it is not:

- (17) *Komo ta pwédeman ése bulá, andá alyí na réyno.* (ZM)
 Since TMA can VB that fly go there LOC kingdom
 'Since he knows how to fly, he goes there to the kingdom.'
 (Forman 1972: 168)
- (18) *Konése hénte ø pwéde éle komé.* (ZM)
 With.these people TMA can he eat
 'With these people he can eat.' (ibid. 170)

In Nubi CA, unmarked stative verbs can generally have both past and non-past reference, as in (19):

- (19) *Ana arufu uwo.* (NB)
 1s know 3s
 'I know him' OR 'I knew him.'

Without further context, non-past reference is considered to be the time of speaking. However, when used in conditional, purposive and certain temporal clauses, the time reference of the verb in these clauses can be that of the verb in the main clause.

- (20) *Kan ana arufu uwo ana bi-asadu uwo.* (NB)
 if/when I know him I FUT-ask 3s
 'If/when I get to know him, I'll ask him.' (ibid.)

In Nagamese, unmarked stative verbs can be interpreted as non-past:

- (21) *Moy gor tat ase.* (NG)
 my house there exists
 ‘My house is there.’ (Sreedhar 1985: 151)

2.2 *Stative verbs with past reference*

While unmarked stative verbs in Atlantic creoles usually have non-past reference, as noted above, they can also refer to the past when this is made clear by the context. Adamson notes that this is the case for Sranan, and Yillah gives the following example in Krio:

- (22) *Olu mɛmba se Agnes nɔ lɛk di os.* (KR)
 Olu think that Agnes NEG like the house
 ‘Olu thought that Agnes did not like the house.’

While variationist studies of JC indicate that stative verbs with past reference are much more likely than non-statives to co-occur with a preverbal marker of anterior tense (Patrick 1992: 343-4), some past-reference stative verbs may be unmarked, as in Sranan and Krio:

- (23) *Ii waan a piis a hais ‘u bai.* (JC)
 ‘He wanted to buy a piece of ice.’

In Guyanese studies there has been some controversy as to whether unmarked statives can have past reference. Bickerton (1975:35-36) argues that *bin* is necessary to mark statives as anterior. However, Gibson (1984) and Jaganauth (1988) have provided counterevidence such as the following:

- (24) *Dem waan tuu maan tu stitch bag.* (GC)
 ‘They wanted two men to stitch bags.’ (Gibson 1984)

In Negerhollands the marking of a verb for past tense is always optional; therefore, there are no constructions in which one typically expects to find a verb with past reference to be unmarked — or marked, for that matter:

- (25) *Fo di ma Kabritabok wet...* (NH)
 Because the man goat knew...
 ‘Because the male goat knew...’ (Stolz 1986: 157)

In Haitian and Dominican CF, unmarked stative verbs such as those in sentences (6) and (7) above could well refer to a past situation, given the appropriate pragmatics.

In the Palenquero CS sentence below, the relative clause triggers deletion of the completive marker *a* according to Green:

- (26) *É jue lo ke uto ø tene de útimo.* (PL)
 he is him who we TMA have of latest
 'He is the one we recently had.' (Friedemann & Patiño Rosselli 1983: 121)

In Papiamentu CS, unmarked stative verbs having past reference do not occur, according to Michel.

In Cape Verdean and Guiné-Bissau CP, unmarked stative verbs can take a past reading given an appropriate context:

- (27) *ma rainha ca credita* (CV)
 but queen NEG believe
 'but the queen didn't believe [that]' (Parsons 1923 cited in Silva 1985: 145)
- (28) *Kil rey ka misti pa ningin kasa-l.* (GB)
 that king NEG want for nobody marry-3s
 'The king didn't want anybody to marry her.' (Kihm 1994: 170)

In the following Angolar sentence, *ta* (the past form of *tha* 'be') in the main clause indicates the time setting of the unmarked stative verb *methe* 'to want' in the subordinate clause, giving it a past reference:

- (29) N *ta* ka pixika, kieke na *methe* mɛ wa. (AN)
 I be-PAST PROG to-fish fish not want eat not
 'I was fishing but the fish didn't bite.'

In Seychellois CF, unmarked stative verbs can refer to the past, given the appropriate context, such as a narrative whose occurrence in the past is made clear at the outset with past markers, but which continues without further tense marking:

- (30) ...*zot ti fin ariv pre avek lakur kot*
 3P ANT COMP arrive near to yard where

ban serpâ i reste...

PL snake SRP stay

‘...they arrived near the yard where the snakes were...’ (Bollée 1977: 161)

In Tok Pisin, stative verbs can refer to the past if accompanied by the appropriate contextual cues:

(31) *Mi save long yu bipo yu kamap tisa.* (TP)

I know LOC you before you become teacher

‘I knew you before you became a teacher.’

In Zamboangueño CS, stative verbs without the preverbal past marker *ya* can refer to the past, depending on context:

(32) *Tódo akél ø kyére ánda saká konése muhér.* (ZM)

All those TMA want go take with.that woman

‘They all wanted to go to take that woman.’ (Forman 1972: 146)

In Nubi CA unmarked stative verbs can also refer to the past:

(33) *Ana arufu de waltumbari.* (NB)

I knew that day-before-yesterday

In Nagamese, verbs are usually inflected to indicate past tense, but stative verbs can be left unmarked:

(34) *Chokri uporte kissa-tu huni-kena, Moa-e tay-ke* (NG)

Girl about story DET hear PAST Moa-NOM she-ACC

bisi bal -pay ø.

much love ?

‘Having heard about the girl, Moa loved her a lot.’

Note that there is a tendency among the younger generation to use the past tense marker *-yse*, i.e. *bal-pa-yse* ‘loved’.

2.3 Non-stative verbs with past reference

Non-stative verbs typically express actions which occur over such a brief period of time that they are considered instantaneous. Adamson states categorically that when non-statives verbs in Sranan have no overt TMA marking, they can only refer to the past:

- (35) A *bay en mama wan moy kado*. (SR)
 3s buy 3s mother a nice present
 'He bought his mother a nice present.'

In Krio action verbs such as *it* 'eat' usually denote the past when they are unmarked:

- (36) *Olu it im petete*. (KR)
 Olu eat his potato
 'Olu ate his potato.'

Also in Jamaican and Guyanese, unmarked non-statives can be expected to have past reference:

- (37) Him *park* him car and we *siddung*. (JC)
 'He parked his car and we sat down.' (Sistren 1986: 48)
 (38) Dem *kom* an dem *res* dem kaafn pan di brij. (GC)
 'They came and they rested their coffin upon the bridge.'
 (Rickford 1987: 241)

In Negerhollands the absence of a tense marker is favored for stative verbs (Sabino 1986: 59), but non-stative verbs can also be unmarked and still have past reference:

- (39) *En mi krii die genade van de heijland nabene my*
 And I got the grace of the savior in my
hart. (NH)
 heart
 'And I received the grace of the Savior in my heart.' (Stolz 1986: 157)

Most common, however, are instances of unmarked verbs with past reference in sentences with two clauses with the same subject, of which the first verb has already been marked for past tense:

- (40) Am *a fang* am, *dra:g* am a hus... (NH)
 He PAST catch him, carry him to house...
 'He caught him, dragged him home...' (Stolz 1986:157)

In Haitian and Dominican CF, there is a noteworthy contrast between stative and non-stative verbs with respect to tense interpretations. Stative

verbs, like those in (6) and (7) above, are interpreted as non-past without regard to the specificity of their objects. However, in Haitian the tense interpretation of non-stative verbs varies with the specificity of their objects: it is past when the object is specific, as in (41a), but non-past when the object is generic, as in (41b):

- (41) a. *Bouki vann chat la.* (HA)
 Bouki sell cat DET
 ‘Bouki sold the cat.’
 b. *Bouki vann chat.* (HA)
 Bouki sell cat
 ‘Bouki sells cats.’

Non-stative verbs are interpreted as past when the object is specific in Dominican CF as well, as in (42a). However, non-stative verbs used without TMA markers do *not* refer to the time of the utterance, even if the object is generic, as we see in (42b):

- (42) a. *Kòkòti vann liv-la.* (DM)
 Kokoti sell book-DET
 ‘Kokoti sold the book.’
 b. **I vann liv.* (DM)
 3s sell book
 ‘He sells books.’

Unlike Haitian, Dominican uses its non-punctual marker *ka* rather than the bare verb to indicate habitual aspect.

In the Palenquero CS sentence below, according to Green, the subordinating adverbial *depue* ‘after’ triggers the zero form, i.e. the deletion of the completive marker *a*:

- (43) *I depué lo k'i ø bota ese basura á*
 and after what.I TMA throw that trash there
 kaposanto... (PL)
 cemetery
 ‘and after I threw that trash there in the cemetery...’
 (Friedemann & Patiño Rosselli 1983: 121)

In Papiamentu CS unmarked non-stative verbs having past reference do not occur, according to Michel.

In Cape Verdean CP unmarked non-stative verbs usually take a past reading:

- (44) *Criad corrê.* (CV)
 servant run
 'The servant ran.' (Parsons 1923 cited in Silva 1985: 145)

The same rule is also true for Guiné-Bissau CP:

- (45) *E bay unde rey.* (GB)
 3P go where king
 'They went to the king's place.' (Kihm 1994: 62)

In Angolar CP, non-stative verbs are like statives in that they do not require preverbal markers to indicate past reference. In the following sentence, unmarked *ba* 'to go' occurred in the context of a personal narration and can only refer to the past:

- (46) *M ba miɔnga.* (AN)
 1s go sea
 'I went to the sea.'

In Seychellois CF the preverbal marker *ti* indicates a time that is past or anterior to that of the context. However, as Corne (1977: 103) points out, once past time has been established in a given situation, *ti* is frequently omitted. This is particularly true of narrative material, and it is often the case in subordinate clauses:

- (47) *Mô ti pe mâze ler mô uar lapli pe vini.* (SC)
 1s ANT PROG eat when 1s see rain PROG come
 'I was eating when I saw the rain coming.' (Corne 1977: 103)

In Tok Pisin, non-stative verbs are assumed to refer to events which occurred in the past, unless marked otherwise by an auxiliary, an adverbial, or by context:

- (48) *Mi kisim kumu long maket.* (TP)
 I get greens LOC market
 'I got greens at the market.'

In Zamboangueño CS, Santoro finds no unmarked non-stative verbs with past reference in Forman (1972).

In Nubi CA, unmarked non-stative verbs are generally interpreted as having past reference:

- (49) *Uwo lebisu gemis da.* (NB)
 3s wear shirt that
 ‘S/he wore that shirt.’

In Nagamese non-stative verbs referring to the past can be left unmarked if their tense reference is clear from the context:

- (50) *Skul-te thak-iya time-te, moy ketiaba modu na-khay.* (NG)
 School-at remain-PROG time in I ever meat NEG-eat
 ‘When I was at school, I never drank (wine).’

However, younger speakers tend to use the past marker in such contexts, i.e. “..na-kha-yse.”

2.4 Non-stative verbs with non-past reference

Unmarked non-stative verbs usually refer to a situation in the past. Adamson states categorically that this is always the case in Sranan, and that the sentence (35) in section 2.3 above cannot have a non-past interpretation.

However, Yillah notes that in Krio, unmarked action verbs such as *mas* ‘step on’ can indeed refer to the present tense when this is clear from the context:

- (51) *Olu mas pen.* (KR)
 Olu step pen
 ‘Olu is stepping on a pen.’

Jamaican has no verbal marker equivalent to the English inflection *-s* for the third person singular of the simple present tense, which usually indicates habitual aspect. In the absence of such a marker (such as *doz* in other Caribbean varieties of CE), non-stative bare verbs with non-past reference are subject to a habitual reading in JC:

- (52) *Him is not a man weh lick.* (JC)
 ‘He is not a man who hits (people).’ (Sistren 1986: 218)

In Guyanese also, unmarked non-statives that have non-past reference typically receive a habitual interpretation:

- (53) *Wenevo yuur having o rilijos fongkshon, dat spirit (GC)*
ov dat porsn ohlwiz kom bak an vizit.
 ‘Whenever you’re having a religious function, the spirit
 of that person always comes back and visits.’ (Rickford 1987: 192)

In Negerhollands, unmarked non-stative verbs — like statives — can have non-past reference:

- (54) *Altit ju praet quaet na bobo sie kop. (NH)*
 Always you speak evil on top his head
 ‘You always speak evil of him.’ (Hesseling 1905: 135)

In Haitian, the tense interpretation of unmarked non-stative verbs is non-past when the object is generic, as in (41b) repeated below:

- (55) *Bouki vann chat. (HA)*
 Bouki sell cat
 ‘Bouki sells cats.’

This contrast with sentence (41a) in 2.3 above is known to Africanists as the factative effect. In the case of Haitian, it has been observed by Damoiseau (1982) and analyzed by Déchaine (1991).

However, as noted in 2.3., in Dominican CF non-stative verbs used without TMA markers do *not* refer to the time of the utterance, even if the object is generic, as in (42b) repeated below:

- (56) **I vann liv. (DM)*
 3S sell book
 ‘He sells books.’

Green points out that in the second Palenquero CS sentence below the lack of a preverbal marker results from the deletion of *ase*, the marker of habitual aspect, rendering non-past tense reference:

- (57) a. *Utere áse ablá por ayá montaña é loma. (PL)*
 you HAB say there mountain is ridge.
 ‘What you call a mountain over there is a ridge.’
 b. *Suto ø ablá montaña.*
 we TMA say mountain
 ‘We say “mountain” here.’ (Friedemann & Patiño 1983: 238)

In Papiamentu CS unmarked non-stative verbs having non-past reference do not occur, according to Michel.

In Cape Verdean CP, according to Silva (1985: 145), the unmarked non-stative constructions which take a non-past reference are usually those “following modals, in conditional and temporal clauses, and in imperatives”. This is the case in the *if*-clause of the conditional construction below:

- (58) *Se bo papia quese portugues es ta sinti*
 if you speak with-them Portuguese they HAB feel
 brigonha. (CV)
 embarrassment
 (Macedo 1979 cited in Silva 1985: 146)

Guiné-Bissau CP appears to have similar restrictions; Kihm (1994: 94) notes that an unmarked verb in a conditional clause can be interpreted as referring to a non-past condition:

- (59) *I kume-u, amanyan no ta padi utru. (GB)*
 1s eat-2s tomorrow 1P FUT conceive other
 ‘(If) he [Lion] eats you, tomorrow we’ll conceive another[child].’
 (Kihm 1994:95)

In Angolar CP, as noted in 2.3 above, unmarked non-stative verbs can only refer to the past and cannot have non-past reference.

In Seychellois CF unmarked non-stative verbs often refer to simple present (60) or habitual (61) actions:

- (60) *Komela mô tir mô marmit lo dife. (SC)*
 now 1s pull 1s pot from fire
 ‘Now I take (am taking) my pot off the stove.’ (Bollée 1977: 54)
- (61) *Mo ser i sât bjê, i kôtâ sâte. (SC)*
 1s sister PM sing well 3s like sing
 ‘My sister sings well; she likes to sing.’ (ibid. 55)

In Tok Pisin, non-stative verbs can be non-past for tense if accompanied by the appropriate contextual cues:

- (62) A: *Nau yu go we? (TP)*
 now 2s go where?
 ‘Where are you going now?’

B: *Mi go long taun.*
 1s go LOC taun
 'I'm going to town.'

In Zamboangueño CS, not only stative verbs (18) but also non-statives verbs like *anda* 'walk' may (63) or may not (64) take the preverbal marker *ta*.

- (63) *Ta anda yo.* (ZM)
 TMA go I
 'I am going.' (Forman 1972: 160)
- (64) *∅ andá alyá na réyno.* (ZM)
 TMA go there to kingdom
 'He goes there to the kingdom.' (ibid. 168)

In Nubi CA, unmarked non-stative verbs generally do not have non-past reference, but in conditional clauses they can have the same time reference as that of the verb in the main clause:

- (65) *Kan ana ayinu uwo,uwo bi-ja ini.* (NB)
 when 1s see 3s 3s FUT-come here
 'When I see him, he'll come here.'

In the sentence below, the unmarked stative verb in the main clause determines a non-past meaning for the entire clause:

- (66) *Kan ana zuru uwo,uwo arufu ana.* (NB)
 if/when 1s visit 3s 3s know 1s
 'When I visit him, he recognizes me.'

In Nagamese unmarked non-stative verbs can have non-past time reference, usually made clear from the context:

- (67) *Moy hoday manso no-khay.* (NG)
 1s always meat NEG-eat
 'I never eat meat.' (Sreedhar 1974:148)

3. Verbs Marked for Anterior or Past Tense

Most of the Atlantic creoles have a verbal marker to indicate a tense anterior to the time in focus; this can correspond to the past, the present perfect or the past

perfect in Western European languages, its interpretation often being sensitive to the distinction between stative and non-stative verbs.

However, in decreolizing varieties, the semantic import of the anterior marker may begin to merge with that of the past tense in the superstrate. Boretzky notes, “It is my impression that wherever an anterior marker is used to express a simple (narrative) past, we are faced with superstrate influence” (quoted in Holm 1988: 151).

3.1 *Stative verbs with past reference*

In Sranan reference to an earlier time is indicated by the preverbal marker *ben*. The exact interpretation of this marker depends on whether the following verb is stative or non-stative verbs. The use of *ben* with statives lends itself to different past tense and aspectual interpretations (e.g. simple past, present perfect, past perfect), whereas its use with non-statives usually corresponds to the past perfect.

- (68) *En granpapa ben lobi en trutru.* (SR)
 3S grandfather PAST love 3S truly
 ‘His grandfather [had] really loved him.’

In Krio the use of the anterior marker *bin* before stative verbs indicates past reference:

- (69) *Olu bin memba di titi wo i bin lek.* (KR)
 Olu ANT remember the girl REL 3S ANT love
 ‘Olu remembered the girl he loved.’

In basilectal JC the preverbal marker of anterior tense is *ben* (with the phonological variants *men*, *wen*, *min* and *en*), but this form is today primarily restricted to rural speakers. The latter also use non-emphatic preverbal *did*, the form that is most common among urban and educated speakers of mesolectal JC:

- (70) *Wa di inglish stuor did niem agen?* (JC)
 what the English store ANT be-called again
 ‘What was the English store called again?’

While there is a quantitative pattern of association between JC stative verbs and the anterior marker, the frequent occurrence of zero marking as in

(23) above makes stronger statements (such as “past statives require *ben/did*”, or “unmarked statives are non-past”) inaccurate for JC.

In Guyanese, Winford (1993: 62-66) notes that stative verbs with past reference are typically marked with anterior *bin* or *did*, but that zero marking is also possible, as in (24) above:

- (71) *Dem bin gat wan lil haus.* (GC)
 3P ANT have a little house.
 ‘They had a little house.’ (Bickerton 1975:35-36)

Ultimately, Winford adopts the explanation of Jaganauth (1988, 1990) that the anterior marker’s typical discourse use is to “background” situations in relation to a reference point, which is the “foreground information.”

In Negerhollands, the preverbal marker *ha* or *a* can be used to indicate past tense with either statives (as below) or non-statives (as in 3.2):

- (72) *Mi a wes en arem verloren Mensch.* (NH)
 I PAST be a poor lost person
 ‘I was a poor, lost person.’ (Stolz 1986:160)

In Haitian CF the tense interpretation of the anterior marker *te* is related to whether the verb it precedes is stative or not. This relationship is summarized (and idealized) as follows (Bickerton 1981):

- a. *te* + stative verb → past (see [73])
 b. *te* + non-stative verb → past-before-past (see [91])

- (73) *Bouki te konn repons lan.* (HA)
 Bouki ANT know answer DET
 ‘Bouki knew the answer.’

However, see discussion of (91) in 3.2 below.

In Dominican CF, the anterior marker *te* is used to indicate past time for stative verbs in a present context:

- (74) *Kókóti te konnèt li.* (DM)
 Kokoti ANT know 3s
 ‘Kokoti knew her.’

In Palenquero CS, the anterior marker is not sensitive to stativity; the morpheme *-ba* attaches postverbally to both stative and non-stative verbs alike. Generally speaking, the anterior marker could be considered a bound

morpheme, although it should be noted that object pronouns can occur between *-ba* and the verb, as in the sentence below:

- (75) *Si majaná nu kelé-lo-ba nu.* (PL)
 yes boys NEG want-it-ANT NEG
 ‘Yes, the boys didn’t want them.’ (Friedemann & Patiño 1983: 267)

Frequently, verbs ending in *-ba* are preceded by the completive aspect marker *a*.

According to Michel, Papiamentu CS does not have an anterior tense (previous to the time in focus in the preceding discourse) but rather a past tense (previous to the time of utterance), for which the preverbal marker is *tabata*. Its use before a stative verb indicates simple past tense:

- (76) *Mi tabata ke kuminda.* (PP)
 I PAST want food
 ‘I wanted food.’

Tabata becomes *taba* before *ta*, the copula and non-past marker, as well as the verb *tin* ‘have’:

- (77) *Mi taba-tin sen.* (PP)
 I PAST-have money.
 ‘I had money.’

In Cape Verdean CP the affixation of the anterior marker *-ba* to a stative verb will give it an immediate simple past reading:

- (78) *Lob’ tem-ba past.* (CV)
 wolf have-ANT food
 ‘The wolf had food.’ (Parsons 1923 cited in Silva 1985: 153)

The same interpretation is found in Guiné-Bissau CP:

- (79) *I ten-ba un kabalu, un baka, i un sancu.* (GB)
 3s have-ANT a horse a cow and a monkey
 ‘There were once a horse, a cow and a monkey.’ (Peck 1988: 228)

In Angolar CP, as indicated above, stative verbs are always unmarked for tense. Thus the anterior (or past) tense marker cannot precede a stative verb.

In Seychellois CF the tense marker *ti* is used with stative verbs to indicate

an anterior time: past time in a present context, or past perfect in a past context:

- (80) *Mô ti kuar i pu al a-pie.* (SC)
 1S ANT believe 3S IRR go on foot
 'I (had) thought he would go on foot.' (Corne 1977: 103)

In Tok Pisin anterior sequence is usually signalled or inferred by contextual cues, adverbials, or by the use of unmarked non-stative verbs; however, the anterior preverbal auxiliary *bin* may be used as well for this purpose.

- (81) *Mi bin save long yu.* (TP)
 I ANT know LOC you
 'I knew you.'

In Zamboangueño CS, the preverbal marker *ya* indicates past rather than anterior tense, and generally has a punctual and perfective meaning, so that it can also indicate completive aspect. The occurrence of *ya* before a stative verb indicates simple past tense, as in (82) below, but note that context can also give unmarked stative verbs past reference, as in (32) above.

- (82) *Ya pwéde le saká konel prinsésa gayót.* (ZM)
 PAST can he seize with-the princess herself
 'He could seize the princess herself.' (Forman 1972: 186)

In Nubi CA the anterior marker *kan* moves the tense of the verb one step backwards in time:

- (83) *Ana kan arufu uwo.* (NB)
 1S ANT know 3S
 'I used to know him.'

In Nagamese past tense is indicated by the morphemes *-ise* (or *-yse* after vowels) and *-ile*, the latter occurring only with locative verbs; the past marker is normally used with both stative and nonstative verbs:

- (84) *Simon Raju-e ki kow-ise itu sob biswas-kor-ise.* (NG)
 Simon Raju-NOM REL say-PAST this all trust-do-PAST
 'Simon believed whatever Raju had to say.'

3.2 Non-stative verbs with (past-before-) past reference

In Sranan non-stative verbs preceded by *ben* usually correspond to the past perfect:

- (85) *Kitty ben gwe bifo Elain doro esde.* (SR)
 Kitty PAST leave before Elain arrive yesterday
 ‘Kitty had left before Elain arrived yesterday.’

In Krio, the anterior marker *bin* has past (rather than past-before-past) reference when used with an non-stative verb:

- (86) *I bin kam.* (KR)
 3S ANT come
 ‘He came.’

Using the completive aspect marker *dɔn* after the anterior marker yields a meaning that corresponds more closely to the English past perfect:

- (87) *I bin dɔn kam.* (KR)
 3S ANT COMP come
 ‘He had come.’

In Jamaican and Guyanese, when the discourse context is already focused on a time before the present, an even earlier point in time can be unambiguously signalled by use of the anterior marker:

- (88) *Sapuoɔz man ben get op an kyatch yu hin de?* (JC)
 Suppose somebody had gotten up and caught you in there?
- (89) *Dem bin gatu get we an kom dis said, lef* (GC)
 They ANT had.to.get away and come over here, leave
di ples an get we, bekaz terabl ting bin
 the place and get away, because terrible things ANT
hapn wid den chiren.
 happened to their children.
 ‘They had to get away and come over here, leave the place and get away, because terrible things had happened to their children.’
 (Bickerton 1975: 36)

In Negerhollands, the preverbal marker *ha* or *a* can be used to indicate past tense with either statives (as in 3.1) or non-statives (as below):

- (90) *Di hon a ne: si fripampi.* (NH)
 The dog PAST take his leash (Stolz 1986: 160)

In Haitian, DeGraff notes that Bickerton (1981) predicts that a non-stative verb marked for anterior tense will be interpreted as past-before-past:

- (91) *Bouki te ale anvan Boukinet vini.* (HA)
 Bouki ANT go before Boukinet come
 'Bouki had left before Boukinet came.'

But Spears (1990) offers a different analysis. In this interesting study of naturally occurring Haitian data, Spears proposes that *te* is an "antiperfect" marker that "...serve[s] to remove a situation from the sphere of the present."

In Dominican CF the marker *te* used with a non-stative verb serves to indicate the anterior time of an action in relation to the context, as in the case of the first verb below:

- (92) *Kókóti te li liv-la avan mwen genyen i.* (DM)
 Kokoti ANT read book-DET before 1s buy 3s
 'Kokoti had read the book before I bought it.'

To indicate anteriority in Palenquero CS, one is not confined to using *-ba*, which also carries the notion of duration, as below:

- (93) *Ané jue-ba lo ke sabe-ba ndrumí loyo.* (PL)
 they be-ANT who know ANT sleep river
 'They were the ones who used to sleep in the stream.'
 (Friedemann & Patiño 1983: 122)

One also has the option of using the past completive marker *a* alone, as below:

- (94) *Konejo a jaya rrancho di tigre bien oganisao.* (PL)
 rabbit TMA find ranch of tiger well organized
 'Rabbit found Tiger's ranch very organized.' (Friedemann & Patiño 1983: 181)

In Papiamentu, *tabata* is used before stative verbs to indicate the past (3.1 above). However, when *tabata* is used with *non-stative* verbs, it indicates past progressive aspect. To indicate the simple past tense of such verbs, the preverbal marker *a* is used:

- (95) Mi *a kumpra* un kas. (PP)
 1S PAST buy a house
 'I bought a house.'

In both Cape Verdean and Guiné-Bissau CP, when the anterior marker *-ba* is affixed to a non-stative verb, the resulting reading is usually that of a past previous to some contextual past already referred to in the discourse:

- (96) *Ele começá ta comê cusa que ele leba-ba.* (CV)
 he start HAB eat thing that he take-ANT
 'He started to eat what he had brought.'
 (Parsons 1923 cited in Silva 1985: 154)

- (97) N *kume-ba* ja oca bu ciga. (GB)
 I eat-ANT already when you arrive
 'I had already eaten when you arrived.' (Peck 1988: 226)

In Angolar CP, the preverbal marker *ta* indicates anterior tense with non-stative verbs; however, as Lorenzino notes, this structure has thus far been attested only when *ta* co-occurs with the progressive or habitual marker *ka*, as in (29) in 2.2. Maurer (1995: 81) concurs: "Unlike the morpheme *tava* in Sãotomense and Principense CP, the Angolar morpheme *ta* does not appear to be able to modify non-stative verbs to indicate anteriority to a past situation. In Angolar, the zero morpheme has that function."

In the case of Seychellois CF non-stative verbs, the anterior marker *ti* must be followed by the completive marker (*f*)*in* to indicate the anterior time of an action in relation to a moment in the past, as below. No simple anterior uses of *ti* (i.e. without the accompanying (*f*)*in*) seem to occur with non-stative verbs:

- (98) *Kâ Zâ ti n gayn* âprizone, (SC)
 when John ANT COMP PASS imprison
Zezi ti vin dân kâtôn Galile.
 Jesus ANT come in area Galilee
 'After John had been imprisoned, Jesus came into Galilee.' (Bollée 1977: 57)

In Tok Pisin, as in the case of past tense, pluperfect tense is usually signalled or inferred by contextual cues, adverbials, or by the use of unmarked non-stative verbs, but the anterior auxiliary *bin* may be used as well for this purpose:

- (99) *Mi (bin) go taun taim yu kam long ples.* (TP)
 I (ANT) go town time you come LOC village
 'I had gone to town when you came home.'

In Zamboangueño CS, non-stative verbs referring to the past must take the preverbal past marker *ya*:

- (100) *Ya andá silá na lamítan.* (ZM)
 PAST go 3P to Lamitan
 'They went to Lamitan.' (Forman 1972: 158)

In Nubi CA unmarked non-stative verbs are usually interpreted as referring to the past, as in (101) below; the anterior marker *kan* moves the tense to past-before-past (102):

- (101) *Ana lebis kabla ita ja.* (NB)
 1s dress before 2s come
 'I dressed before you came.'
- (102) *Kan ana lebis kabla ita ja.* (NB)
 ANT 1s dress before 2s come
 'I was/had [already] dressed before you came.'

In Nagamese the morpheme *-ise* or *-yse* normally marks non-stative verbs as having past reference:

- (103) *Moa ketiya Calcutta ja-yse?* (NG)
 Moa when Calcutta go-PAST
 'When did Moa go to Calcutta?'

4. Conclusions

Our findings can be summarized in the following table, in which (+) indicates that a structure is attested, (0) indicates that its absence is attested, (?) indicates that its presence is unknown, and R indicates that its presence is rare.

Table 1. *Stativity and tense reference in various creole languages*

	<i>Atlantic</i>				<i>Non-Atlantic</i>	
	SKJG	NHD	LP	CEA	YZT	BM
<i>Unmarked verbs</i>						
<i>Statives with non-past ref.</i>	++++	+++	++	+++	+++	++
<i>Statives with past ref.</i>	++++	+++	+0	+++	++R	++
<i>Non-statives with past ref.</i>	++++	+++	+0	+++	+0+	+0
<i>Non-stat. w/ non-past ref..</i>	0+++	+++	+0	++0	++R	++
<i>Anterior (or past) marker</i>						
<i>Statives with past ref.</i>	++++	+++	++	++0	+++	++
<i>Non-stative with (past-before) past ref.</i>	++++	+++	++	++0	+++	++
Creole language name abbreviations:						
S = Sranan CE	L = Palenquero CS		Y = Seychellois CF			
K = Krio CE	P = Papiamentu CS/P		Z = Zamboangueno CS			
J = Jamaican CE			T = Tok Pisin C/PE			
G = Guyanese CE	C = Cape Verdean CP					
	E = Guiné-Bissau CP		B = Nubi CA			
N = Negerhollands	A = Angolar CP		M = Nagamese P/CAs			
H = Haitian CF						
D = Dominican CF						

The following general pattern can be seen in the above table: in most of the creoles surveyed, whether Atlantic or non-Atlantic, both stative and non-stative verbs can have either past or non-past reference when unmarked, and past (or past-before-past) reference when marked for anterior or past tense.

The two creoles with salient differences from the others are Papiamentu CS and Angolar CP. In Papiamentu, unmarked verbs (whether stative or non-stative) cannot refer to the past, and non-stative verbs can refer to the non-past only if they are marked for the present with *ta*; i.e. unmarked non-stative verbs with non-past reference do not occur. This is similar to the system of obligatory tense marking in the European languages from which Papiamentu derived most of its lexicon (Spanish and Portuguese), which (along with Dutch) continued to influence the creole via the native speakers of these languages who were also native speakers of Papiamentu, who seem likely to have influenced the Papiamentu of monolinguals. The hypothesis that Papiamentu's syntax is among the most Europeanized of the Atlantic creoles, having never undergone the morphological decreolization characteristic of many English-based Caribbean creoles, is borne out by other features of Papiamentu syntax,

such as the copula. This follows the pattern of European languages in which a single-form copula is obligatory in those syntactic environments which require the absence of a copula or a completely different copula in most Atlantic creoles (Holm et al. 1999).

Angolar resembles Papiamentu in that unmarked non-stative verbs cannot have non-past reference; however, in Angolar these verbs have past reference, whereas in Papiamentu they do not occur. Furthermore, in Angolar, stative verbs are always unmarked for tense and cannot take a past or anterior marker, while non-stative verbs can only take an anterior marker when it co-occurs with a progressive aspect marker. The origin of this pattern remains unknown.

These differences aside, the similarity between the remaining creoles, whether Atlantic or non-Atlantic, is remarkable. Although the pattern is unlike that of Western European languages, which require tense marking, it is still possible that this pattern is the result of a semantic universal regarding the nature of stativity. A stative verb by definition refers to a state of affairs, which is likely to be open-ended and to extend into the present, whereas a non-stative verb refers to an action whose duration may be practically instantaneous, i.e. not open-ended and thus more likely not to extend into the present.

The creole Portuguese of Guiné-Bissau is typical of most of the creole languages in this survey: as Suzuki notes, unmarked stative verbs in GB are generally interpreted as referring to present states or states relevant at the point in time in focus in the discourse, whereas unmarked non-stative verbs are interpreted as referring to past situations (Peck 1988: 181). The difference between statives and non-statives seems to be attributable to the fact that the states described by unmarked stative verbs can frame a reference point in time, whereas the events or perfective situations described by unmarked non-stative verbs cannot. Thus, when a non-stative verb is combined with a marker of imperfective (or durative or non-punctual) aspect — e.g. progressive *na* or habitual *ta* — it is generally interpreted as referring to a situation occurring or relevant in the present or at the point in time in focus in the discourse (Peck 1988: 288).

Noticing this tendency in Guyanese CE, Bickerton concluded that the unmarked verb “...signifies past with nonstatives and nonpast with statives” (1979: 309). Although Gibson (1984) demonstrated that this was not always the case in Guyanese, Bickerton’s more general claim about tense marking and stativity (“...there is strong evidence for supposing that its Guyanese

function is common if not universal in Anglo-Creoles generally” [ibid.]) has been widely accepted and extended to all Atlantic and even non-Atlantic creoles.

Our data demonstrate that while Bickerton’s claim about the relationship between stativity and tense reference may reveal an interesting statistical tendency reflecting the inherent meaning of stativity, his claim has no predictive value in any of the Atlantic or non-Atlantic creoles surveyed in this study.

Abbreviations

A = Arabic; ACC = accusative; AN = Angolar CP; ANT = anterior marker; As = Assamese; C = Creole; COMP = completive marker; CV = Cape Verdean CP; D = Dutch; DET = determiner; DEM = demonstrative; DM = Dominican CF; E = English; F = French; FUT = future marker; GB = Guiné-Bissau CP; GC = Guyanese CE; HA = Haitian CF; HAB = habitual aspect marker; IRR = irrealis marker; IMP PN = impersonal pronoun; JC = Jamaican CE; KR = Krio CE; LOC = locative; NB = Nubi CA; NG = Nagamese CAs; NH = Negerhollands CD; NOM = nominative; P = Portuguese; PASS = passive; PAST = past tense marker; PL = Palenquero; PM = predicate marker; PP = Papiamentu; PRES = present tense; PROG = progressive aspect marker; PT = past participial inflection; REL = relative; S = Spanish; SC = Seychellois CF; SR = Sranan CE; SRP = subject referencing pronoun; TMA = tense/mood/aspect marker; TP = Tok Pisin P/CE; VB = verbalizing morpheme; ZM = Zamboangueno CS; 1S = first person singular pronoun or possessive, 2P = second person plural, etc.

Note

- * An earlier version of this paper was presented at the conference of the Society for Pidgin and Creole Linguistics held in San Diego in January, 1996. This version has profited from the discussion there, and from the suggestions of John McWhorter and an anonymous reviewer. However, responsibility for any remaining shortcomings is solely our own.

References

- Agheysi, R. 1971. *West African Pidgin English: Simplification and Simplicity*. Stanford University PhD dissertation.
- Bailey, B.L. 1966. *Jamaican Creole Syntax*. Cambridge: Cambridge University Press.
- Bickerton, D. 1975. *Dynamics of a Creole System*. Cambridge: Cambridge University Press.

- Bickerton, D. 1979. "The status of *bin* in the Atlantic Creoles". In *Readings in Creole Studies*, I. Hancock, E. Polomé, M. Goodman and B. Heine (eds.) 309-14. Ghent: E. Story-Scientia.
- Bickerton, D. 1981. *Roots of Language*. Ann Arbor: Karoma.
- Bollée, A. 1977. *Le Créole Français des Seychelles*. Tübingen: Max Niemeyer.
- Comrie, B. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- Corne, C. 1977. *Seychelles Creole Grammar: Elements for Indian Ocean Proto-Creole Reconstruction*. Tübingen: Günter Narr.
- Damoiseau, R. 1982. *Étude de Syntaxe Comparée*. Université d'État d'Haiti.
- Déchainé, R-M. 1991. "Bare sentences". In *Proceedings from the First Conference on Semantics and Linguistics Theory 10*, S. Moore & A.Z. Wyner (eds), 31-50. Ithaca: Cornell University.
- Forman, M. 1972. *Zamboangueno with Grammatical Analysis*. Cornell University PhD dissertation.
- Friedemann, N.S. de and Rosselli, C.P. 1983. *Lengua y Sociedad en el Palenque de San Basilio*. Bogotá: Instituto Caro y Cuervo.
- Gibson, K. 1984. "Evidence against an anterior tense system in Guyanese and Jamaican Creoles". In *York Papers in Linguistics* (Papers from the York Creole Conference), M. Sebba and L. Todd (eds.), 123-129. Department of Language, University of York.
- Hesseling, D.C. 1905. *Het Negerhollands der Deense Antillen*. Leiden: A.W. Sijthoff.
- Holm, J. 1988. *Pidgins and Creoles* (Vol. I). Cambridge: Cambridge University Press.
- Holm, J. (ed.) forthcoming. *Comparative creole syntax*, J. Holm and P. Patrick, (eds). London: U. of Westminster Press.
- Holm, J. et al. 1999. Copula patterns in Atlantic and non-Atlantic creoles. In J.R. Rickford and S. Romaine, eds. *Creole genesis, Attitudes and Discourse: Studies celebrating Charlene Sato*. Amsterdam, New York: John Benjamins, pp. 97-119.
- Jaganauth, D. 1988. "Relative time reference in Guyanese Creole: some problems for sentence level analyses". Paper presented to the Society for Caribbean Linguistics, Bahamas.
- Jaganauth, D. 1990. "Creole time-reference markers in discourse". Paper presented to the Society for Caribbean Linguistics, Belize.
- Kihm, A. 1994. *Kriyol Syntax: the Portuguese-based Creole Language of Guinéa-Bissau*. Amsterdam: John Benjamins.
- Macedo, D.P. 1979. *A Linguistic Approach to the Capeverdean Language*. Boston University D. Ed. dissertation. Ann Arbor: University Microfilms.
- Maurer, P. 1995. *L'Angolar: Un Créole Afro-portugais Parle à São Tomé*. Hamburg: Buske Verlag.
- Parsons, E.C. 1923. *Folk-lore from the Cape Verdean Islands* (2 volumes). Cambridge, Mass. and New York: American Folklore Society.
- Patrick, P.L. 1992. *Linguistic Variation in Urban Jamaican Creole: A Sociolinguistic Study of Kingston, Jamaica*. University of Pennsylvania PhD dissertation.
- Peck, S.M., Jr. 1988. *Tense, Aspect and Mood in Guinéa-Casamance Portuguese Creole*. Ph.D. dissertation, University of California, Los Angeles PhD dissertation. Ann Arbor: University Microfilms.

- Rickford, J.R. 1987. *Dimensions of a Creole Continuum*. Stanford: Stanford University Press.
- Sabino, R. 1986. "Another step towards a characterization of the Negerhollands tense and aspect particles". *Amsterdam Creole Studies* 9:47-71.
- Sankoff, G. 1990. "The grammaticalization of tense and aspect in Tok Pisin and Sranan". *Language Variation and Change* 2: 295-312.
- Silva, I. 1985. *Variation and Change in the Verbal System of Capeverdean Crioulo*. Georgetown University PhD dissertation. Ann Arbor: University Microfilms.
- Sistren (ed) (with H.F. Smith) 1986. *Lionheart Gal: Life Stories of Jamaican Women*. London: The Women's Press, Ltd.
- Spears, A. 1990. "Tense, mood and aspect in the Haitian Creole preverbal marker system". In *Pidgin and Creole Tense-Mood-Aspect Systems.*, J. Singler (ed), 119-42. Amsterdam: John Benjamins.
- Sreedhar, M.V. 1974. *Naga Pidgin: A Sociolinguistic Study of Inter-lingual Communication Patterns in Nagaland* (Occasional Monograph Series 26). Mysore: Central Institute of Indian Languages.
- Sreedhar, M.V. 1985. *Standardized Grammar of Naga Pidgin* (Occasional Monograph Series 26). Mysore: Central Institute of Indian Languages.
- Stolz, T. 1986. *Gibt es das kreolische Sprachwandelmodell? Vergleichende Grammatik des Negerholländischen*. Frankfurt: Peter Lang.
- Winford, D. 1993. *Predication in Caribbean English Creoles*. Amsterdam: John Benjamins.

Are Creole Languages “Perfect” Languages?

Alain Kihm

If humans could communicate by telepathy, there would be no need for a phonological component, at least for the purposes of communication, and the same extends to the use of language generally. These requirements might turn out to be critical factors in determining the inner nature of C_{HL} in some deep sense, or they might turn out to be “extraneous” to it, inducing departures from “perfection” that are satisfied in an optimal way. The latter possibility is not to be discounted. (Chomsky 1995:221)

1. Introduction

Depending on one’s theoretical persuasions or current mood, the above quotation from Chomsky’s much-cited Chapter 4 may strike one as a bit cranky or imbued with deep significance. Actually, it raises a real issue concerning the relationship between the inner and outer forms of language, and the type of determinism that unites them. In particular, interesting consequences follow when one starts thinking over the key concept “perfection” that it contains. This article is devoted to exploring this concept and its implications for creole studies.

First, I will examine in what sense human language(s) might be perfect, and in what sense they are not. Then, I will consider the relevance of the “perfection” concept to creole languages given the current theories about their genesis. Finally, I will propose some ideas about the partially contradictory observations that creoles are a) imperfect like any other languages; and b) less imperfect than other languages, a character that is often expressed as their being (relatively) “simple”. In sum, I would like the present work to be an examination of how creolization hypotheses, especially the Language Bioprogram Hypothesis, fare with respect to recent developments in linguistic theory

such as Minimalism. These developments raise crucial challenges in imposing increasingly strict and precise constraints upon the shape of Universal Grammar, a good example being Kayne's (1994) Universal Base Hypothesis on which more will be said below (also see Zwart 1997). I will therefore explore the assumptions that, in my opinion, any universalist theory of creole genesis must make and the consequences they have. The present work also aims to be a contribution to the debate between competing hypotheses in the field of creole studies, in particular Language Bioprogram vs. Relexification. In fact, I have nothing really new to say. My only originality, if any, is that of entering the problem from a hitherto untrodden perspective, at least to the best of my knowledge.

2. Why are human languages not perfect?

What does it mean to predicate such an attribute as "perfection" of language? Chomsky (1995) makes it clear (e.g. p. 9) that "perfection" is related to the sister notions of uniformity, simplicity, and non-redundancy, and it must be distinguished carefully from another concept that is often, but wrongly, associated with it, namely that of "adaptation". That language as a biological endowment is adapted to the survival of the human species seems plain enough, since we have been here and talking for perhaps a million years. This fact does not imply that language should be uniform and simple in any relevant sense. On the contrary, biological adaptation is generally seen to be an extremely complex, untidy phenomenon. Actually, complexity and variation are certainly more adaptive than are their contraries (see Weinreich, Labov, & Herzog 1968 for a linguistic argument to that effect).¹

Language, however, can also be viewed as a biological *computational* system, from which all other computational systems such as mathematics, formal logic, etc. are possibly derived as specialized artificial applications. It seems therefore reasonable to assume that language, i.e. internalized language or grammar (see Chomsky 1986), is submitted to the same specifications of simplicity, uniformity, and non-redundancy as are these other computational systems. To put it differently, there may be a level at which it makes sense to view language as a potentially perfect system that departs from perfection because it can only manifest itself by being interfaced with other parts of the organism for which perfection requirements have no relevance. Those other

parts consist minimally in the sensorimotor systems (regulating vocal productions or gestures) and the systems of thought, and the question may be formulated in terms of the extent to which language is well designed to meet the "legibility" conditions imposed by these systems (see Chomsky 1998).² The strongest thesis, then, is that "language is an optimal solution to legibility conditions" (Chomsky 1998:9), where "optimal" may be understood as "as perfect as possible".

It is also important to emphasize that the biological and computational dimensions are not separate entities, but different levels of the same, ultimately neuronal reality. This point is related to a recent discussion by Li (1997) of the appropriateness of redundancy removal in linguistic theory, when all biological systems show themselves to be highly redundant. Li argues that the elimination of redundancies is a legitimate and desirable task insofar as language is treated as a black box, the inner cogs and wheels of which are hidden from us. These inner mechanisms (i.e. neuron assemblies), in contrast, may well be redundant in their functioning. Their discovery, however, would not invalidate the search for a non-redundant theory of language, because the level at which language is a computational system is just as real in the architecture of the mind-brain as the level at which it resolves itself into bundles of firing neurons.

The question, then, is whether departures from perfection are random or definable. Of course, we have to assume the latter, since accepting the former would terminate all inquiry. Given what little is known of the external levels of sound and thought, especially the latter, it is probably impossible to say anything meaningful about how they are related "through" language. What we may do with more safety, however, is to inquire into the relation between the internal interfaces with these levels, about which substantial theories are at least available. The language-thought interface we shall call "(narrow) syntax" or following Brody (1995), "Lexico-Logical Form" (LLF). Syntax is then an arrangement of syntactic objects usually called categories (see Collins 1997 for a clear exposition). I will further adopt Kayne's (1994) Universal Base Hypothesis that categories are uniformly ordered according to a Linear Correspondence Axiom (LCA) that derives succession from the structural relation of asymmetric c-command. The language-sound interface, on the other hand, is an arrangement of forms which may be called the morphophonological form (MF: see Halle & Marantz 1993; Marantz 1997). Other hypotheses are of course possible, and mismatches will vary accordingly in

tokens and in types. But no theory of language, I believe, no matter how “surfacy”, can seriously maintain that what we say or hear is exactly what we mean or understand.

From these assumptions, two lines of inquiry follow. One consists in asking whether MF and syntax are well designed in their function as interfaces with the external systems of sensorimotoricity and thought respectively (Chomsky’s 1998 strong thesis). This is not the line I will pursue, as it far exceeds my purpose and capacities. I will tackle a more limited issue, also entailed by the general position of the problem and more directly relevant to the question posed in the title of this paper, namely: assuming the interfaces to be in themselves optimal, is their mutual relation also optimal and, insofar as it is not, what makes it less than optimal? That is to say, the present work aims to be an inquiry into language complexity and whether creole languages, as a consequence of their genesis, are or should be significantly less complex (i.e. endowed with a “better” syntax-MF relation) than non-creole languages.

What is an optimal relation between the interfaces syntax and MF?³ Isomorphism is optimal, in the sense that the complex syntactic object $\{SO_1, SO_2, \dots\}$ maps bijectively onto the complex morphophonological object $\{\phi_1, \phi_2, \dots\}$. To put it differently, grammar in the optimal case would be a characteristic function over the set syntax-MF such that, taking SO_i as arguments and ϕ_i as their values, assigning a value to an argument would always yield a truth value — and conversely, taking ϕ_i as arguments and SO_i as their values.

Of course this is not the usual picture. What we generally find is partial mismatch between syntax and MF, meaning that it is not always possible to pair them through a characteristic function, either because some SO has no value at MF, or because some ϕ does not correspond to an SO. *A priori*, the origin of the mismatch may reside with either one or both. That is to say, there are three possibilities: a) MF distorts syntactic structure; b) syntax distorts MF; c) syntax and MF obey different principles. We can rule out b) if we want to keep to the general hypothesis that syntax is what is truly universal in language, whereas MF has many parochial departments. Now a) and c) will combine given a derivational view of grammar involving three steps such that: i) a syntactic structure is built; ii) it is converted into an MF structure through the function Spell Out; iii) the MF structure is interpreted by the phonological component in the strict sense. This is the view implicit in Chomsky’s quotation. According to it, the mismatch between the surface and the base (what is heard and what is meant, roughly speaking) is mainly due to step ii) above.

Morphology destroys optimality in three ways.

Firstly, it is the level where the *arbitraire du signe* becomes manifest. To understand fully what this means, we have to make a distinction between two different objects that are often confused, viz. the lexicon and the vocabulary (see Halle & Marantz 1993; Marantz 1997, 1999). The lexicon is the repertoire of elements used as building blocks for syntactic objects. It is linked with syntax on the one side and with the conceptual-intentional (or thought) system on the other side, and the items it consists in may be viewed as semantic features bundled up in specific ways to form lexical items or *roots*. The vocabulary, on the other hand, is a repertoire of *words*, and it is linked with MF and the articulatory-perceptual system. In this framework, lexical insertion, as it is traditionally called, is in fact vocabulary insertion, and it is a late process, occurring in MF.

We also have to take seriously the idea that there is one human language realized under various guises. Given this and the foregoing, the *arbitraire du signe* takes two forms. On the one hand, there is the fact that applying the function Spell Out to a root will not always return the same word, and that the outcome is arbitrary. This is the traditional meaning of the Saussurean notion, expressing the observation that the root TREE is realized here as *tree*, there as *arbre*, etc. — assuming all human beings to share the same or compatible representations of what a “tree” is, a plausible assumption, it seems.⁴ Put differently, Spell Out as a function from the lexicon to the vocabulary is not one-place but two-place, i.e. taking two arguments, a root and a particular language name, so that, e.g., Spell Out (TREE, English) = *tree*.⁵ Another way that signs are arbitrary is partially related to the second and third points below: the mapping of roots onto words is not necessarily one-to-one as with the example just given. For instance, assuming the entirely putative complex feature bundle {GO TAKE COME}, we see that it can be mapped onto the one word *fetch* in English, whereas it is mapped onto the three-term serial verb *am d ap* “go hold come” in the Papuan language Kalam (see Foley 1986).⁶ In other words, *fetch* is a vocabulary item, but it is not a lexical item.

Secondly, morphology induces displacements, i.e. elements are pronounced in different positions from those they are interpreted in because of the necessity of overtly checking functional features. For instance, we find the following well-known mismatch in many languages:

- | | | |
|-----|--------|-----------------------------|
| (1) | Syntax | $[_{TP} T... [_{VP} V...$ |
| | MF | $[_{TP} V-T... [_{VP} t...$ |

In syntax, the functional category Tense dominates the verb root and, according to the LCA, precedes it; at MF, in contrast, V has raised and adjoined to T, so the tense morpheme now follows the verb base or is not ordered with respect to it. As a result, the syntactic object {T V} gets mapped onto what is ultimately one word, such as French *écrivait* ‘s/he wrote’, where the Tense morpheme /-ε(t)/ follows the stem /ekriv-/, or Arabic *kataba* ‘he wrote’, where the Tense morpheme /a-a/ neither precedes nor follows the root $\sqrt{\text{KTB}}$, to take two languages where it seems fairly uncontroversial that the verb moves to T. Notice that MF displacement ought to be distinguished from scope movement whereby an operator is brought to a position where it can bind a variable, WH-movement being an example. Indeed, in “Who did you meet yesterday?”, interpreted as “for which x , x a person, did you meet x ?”, there is every indication that *who* does appear at MF in its syntactic or LLF position, i.e. dominating the whole proposition. This implies that elements such as *who* should be conceived of as discontinuous items generated in two positions simultaneously because of their dual nature as operators and arguments (or place-holders for arguments). In this way, pronouncing the first occurrence or copy as in English or the second one as in colloquial French *Tu as rencontré qui hier?* (the so-called “in-situ strategy”) does not induce any real distortion between syntax and MF as schematized below:

- (2) Syntax: $[_{\text{CP}} \text{WH} \dots [_{\text{VP}} \dots \text{WH}]$
 MF: $[_{\text{CP}} \text{WH} \dots [_{\text{VP}} \dots \text{WH}] = \textit{Qui tu as rencontré hier?}$
 $[_{\text{CP}} \text{WH} \dots [_{\text{VP}} \dots \text{WH}] = \textit{Tu as rencontré qui hier?}$

Thirdly, morphological operations are a source for multicategoriality (see Solà 1996), i.e. the property that words have of spelling out several categories corresponding to different functional heads in addition to one or several lexical heads. This type of one-to-many mapping, detrimental to isomorphism, has to be distinguished from the one embodied in *fetch* that was mentioned above. With *fetch*, several roots are mapped onto one word through the function Spell Out operating between the two levels. Consequently, category is uniform before and after Spell Out.⁷ Consider now a multicategorial word such as Arabic *katabtu* ‘I wrote’, consisting in the root $\sqrt{\text{KTB}}$, the Tense morpheme /a-a/, and the Agreement (person-number-gender, a.k.a. ϕ -features) morpheme /-tu/ ‘1st person singular’.⁸ It is the result of Spell Out operating between syntax and MF. Isomorphism is destroyed in two ways in this case. First, ordering is reversed or suppressed as shown in (3):

- (3) Syntax: [_{TP} Past [_{AgRP} 1SG [_{VP} √KTB
 MF: [_{V-T-Agr} katabtu

Secondly, insofar as it is a vocabulary item at MF, *katabtu* is an inflected verb. In syntax, on the other hand, it is a predicate with a subject. From a semantic point of view, there is thus no uniform correspondence between the types of the “same” element at different levels.⁹ According to Pustejovsky (1995), such type shiftings conducive to generalized polysemy are desirable since they allow for the creative use of language. Language would therefore be less perfect without them. Notice, however, that we are here using the adaptation sense of “perfect”: a polymorphic language is better suited to the purposes human language has to fulfill than a monomorphic one would be (see Pustejovsky 1995). As mentioned, this is a different issue from that of computational perfection or optimality.

Of course, the mismatches schematized in (1) and (3) are related, as they both result from the morphological operations “merger” (Halle & Marantz 1993). They should not be confounded, however, since not all morphologically-driven displacements correlate with morphological merger. This issue will be taken up below.

Finally, a special combination of multicategoriality with the *arbitraire du signe* gives us suppletion as standardly exemplified by the English pair *go-went*. Suppletion is traditionally considered a marginal phenomenon affecting only “light” elements with no or little referential content. It is not expected with items like *table* or *write*, so the plural of the former would be, e.g., *wabe*. This is probably correct, but there are borderline cases, nevertheless, as shown by the Neo-Caledonian language Xârâcùù where one finds such pairs as *kâmûrû* ‘man’ vs. *pââdo* ‘men’, or *çê* ‘woman’ vs. *paê* ‘women’ (Moyse-Faurie 1995). Hardly less exotic is Ancient Greek *horaô* ‘I see’ vs. *eidon* ‘I saw’.¹⁰ What suppletion brings about is extreme dysmorphism of the type exemplified in (3) since, contrary to *katabtu*, *went* includes no subparts that could be mapped to elements of the syntactic object {T GO} it spells out.

To summarize, some of the factors that induce departures from optimality may be identified as having to do with the intermapping of syntax with MF, while another part concerns the relation of the outer edges of these interface components, viz. the lexicon and the vocabulary. More generally, both types of departures may be seen as resulting from the same process of plugging a computational system into intrinsically complex levels: sounds (or gestures) on the one hand, concepts and representations on the other hand.

From an evolutionary perspective, it is quite conceivable that the latter level (perhaps the former as well) has to do with what Bickerton (1990, 1998) calls the “protolanguage”. I will return to this issue. As far as the syntax-MF linking is concerned, the existence of distorting morphophonological processes such as have been exemplified is probably due ultimately to the intrinsic fragility and mutability of sound as a medium.¹¹ That is to say, even though syntax-MF intermapping appears to be internal to grammar, its departures from optimality may be considered repercussions from extraneous requirements, that would not occur if these requirements were absent. Were it not for the need to speak it, language would be closer to perfection. Surprising, or even scandalous as this idea may sound, I think it has much to commend it.¹²

3. Why should creole languages be perfect?

The answer to this question must depend on one’s view of creolization. If creolization is conceived of as a more or less special case of language change, then there is no reason why creole languages should be perfect or more optimal than any other natural languages. I am referring here to all theories that assign a crucial role to the notions of “contact” and “second language acquisition”, whether or not they maximize the part played by the substrate (see, e.g., Lefebvre 1986; Thomason & Kaufman 1988). Only if we accept the Language Bioprogram Hypothesis (henceforth LBH) as a plausible theory of creole genesis does the question become meaningful and even unescapable.¹³ Only then, as will be explained more fully below, does it follow that creoles should be optimal or nearly optimal languages. I introduce this qualification in order to dispel a possible objection, namely that most creole languages have existed for at least three centuries during which they have incurred all the accidents of history and may thus have steered away from perfection. It seems, however, that major morphophonological processes distorting the original design in the direction of the dysmorphisms exemplified in the preceding section would have taken more time than a mere three centuries (in which light it is important that many creoles have existed for even less time than this). If creole languages were optimal at birth, then, they should still be recognizably so, in the sense that the overall degree of syntax-MF isomorphism in their grammars should be greater than in non-creole languages. That is, of course, unless we want to assume that the original design was distorted almost as soon

as it was realized by virtue of being realized — an interesting idea which I shelve for the moment.

Why, then, should creole languages born according to the LBH be perfect (i.e. optimal), at least in principle, i.e. abstracting from what they may have incurred in later years? That follows directly from the hypothesis. Indeed, creole languages in this theory are assumed to represent the actualization in the child’s brain not of some preexisting language, i.e. of some parameterization of UG, but of the language bioprogram itself (see Bickerton 1984). Clearly, the language bioprogram cannot be a distinct object from Chomsky’s C_{HL} , i.e. the computational system of human language. That is to say, it should be possible to substitute one term for the other in the epigraph *salva veritate*. Therefore, the language bioprogram should be optimal to the extent and in the same sense that C_{HL} is optimal.

There are two ways that this conclusion can be rejected. One is simply to deny that the whole issue makes sense by considering the distorting factors to be part of C_{HL} and not the consequence of extraneous requirements. Chomsky’s formulation allows for this possibility, but he clearly views it as theoretically less interesting as well as probably less plausible empirically.¹⁴ The other way is to stress the “bio” prefix, that is to claim that the bioprogram is in fact distinct from C_{HL} in that the latter only includes the computational aspect of human language, whereas the former is the neuronal instantiation of language as a biological organ. In that sense, the bioprogram could only be said to be perfect insofar as it is optimally adapted to the continuing survival of the human species. Perhaps this is the right way to interpret the notion.

Yet, such an interpretation is totally irrelevant to the problem at hand. Whatever its ultimate reality, the bioprogram has to encompass all the aspects of language touched upon in the preceding section. Of these aspects, the syntax-semantics lack of isomorphism due to the polymorphic character of language may have adaptive value. It is hard to imagine, however, how the morphologically induced departures from isomorphism schematized in (1) and (3) might fit into a Neo-Darwinian paradigm.¹⁵ It is therefore impossible to divorce the bioprogram from the computational dimension of human language. At the very least, the bioprogram must include this dimension as a proper subset. There is no conceptual or real necessity, in contrast, that it should include those “extraneous” factors that distort perfection on the one hand, and apparently have no biological adaptive function on the other hand. In other words, the alternative is this: either the bioprogram is fully coexten-

sive with C_{HL} , in which case it does not include the “extraneous” factors and is perfect in that sense; or it names the set consisting in C_{HL} plus what is biologically adaptive in language, and it still does not include the “extraneous” factors. Since it makes no difference for my purpose, I will choose the first member of the alternative.¹⁶

Under this view, the bioprogram consists of a syntax and a formal semantics, that is to say it may be equated with the narrow syntax or LLF component of grammar, apparently excluding the MF component, because of its more direct dependence on extraneous factors and its irreducible parochiality.¹⁷ It also excludes, albeit in a different sense to which I will return, the external “general” semantic component. In this way, the bioprogram is but a limited part of language as a whole. We may speculate, as Bickerton (1998) indeed does, that it expresses the small but highly consequential mutation that separated *Homo sapiens* from its forefathers, an idea related to the already mentioned “protolanguage” notion.

An actual language mirroring the bioprogram will thus be optimal, except to the extent that it has to be spoken. Since all MF processes (except *Spell Out per se*) are excluded by hypothesis, words in such a language are direct expressions of lexical items, occurring in the same temporal order that these items do in the syntactic structures they build, so the only, unavoidable departure from optimality is the *arbitraire du signe* in the simplest sense. What this points to is complete isomorphism of the vocabulary and the lexicon. In particular, items such as *fetch* are not expected in a perfect language, given a plausible notion of possible roots as semantic feature bundles — corresponding, e.g., to Hale & Keyser’s (1993) Lexical Relational Structures.

In a way, all this is but a reformulation of the hackneyed remark that creole languages have no or next to no morphology. To a large extent, the remark expresses a true observation. Fully spelling out the notions involved, as I tried to do, is then of interest in showing that absence of morphology is also an intrinsic consequence of a given creolization theory, at least as much as a contingent fact that the theory wants to explain. On the other hand, the present argument should not be mistaken for the also common notion that creole languages are somehow “semantically transparent” (see Seuren & Wekker 1986). If formal semantics, i.e. LLF, is intended, then creole languages are indeed transparent insofar as they are optimal. But optimality in that sense does not imply transparency from the viewpoint of semantics more

generally understood. For instance, it is well-known that *window* is a polysemous item denoting either an aperture (e.g., “John crawled through the window”) or an object (“Mary broke the window”).¹⁸ Should a semantically transparent language show two different words here? Unless it is demonstrated that sense can be made of this question, which I doubt, “semantic transparency” remains too vague a notion to be of much use, it seems to me.¹⁹

4. Are creole languages perfect?

Given the foregoing argument, the question has to be broken down into three factual queries: i) Do creole grammars include morphologically driven displacements (overt movement)? ii) Do they show mergers leading to multicategoriality? iii) Do they include suppletions?

The answer to i) and ii), which we may treat jointly, is definitely “yes”. Take a Haitian DP such as *chat-la* ‘the cat’ (see Joseph 1988; Lefebvre & Massam 1988). According to our assumptions, it can be assigned the following syntactic structure:²⁰

$$(4) \quad [_{DP} LA [_{NumP} \emptyset [_{NP} chat$$

The order in (4) is exhaustively determined by general phrase structure principles (functional categories asymmetrically c-command the lexical categories they modify), universal feature hierarchy (determination dominates number), and the LCA (heads precede their complements). Each category is uniquely expressed by a lexical item, including $[_{NumP} \emptyset]$, the unmarked member of the binary contrast $[\pm plural]$.²¹ Is (4) an optimal expression? Not quite, because it does not show that the definiteness and number features are matched on the head noun. One way to represent this fact is to copy the relevant features of this head in positions where they can be checked by the corresponding features of the functional projections D and Num, i.e. in the specifiers of these projections. The full syntactic structure is therefore (5), where FF_N designates the relevant functional features of the head noun:

$$(5) \quad [_{DP} FF_N [_{D^o} LA] [_{NumP} FF_N [_{Num^o} \emptyset] [_{NP} [_{N^o} chat$$

In this structure, we have the representation of LF or covert movement, involving only features.²² Since functional features have no phonological form (or are deleted when checked), the straightforward spelling out of (5)

would yield the ungrammatical Haitian sequence **la chat*. In order to derive *chat-la* ‘the cat’ we therefore have to assume that not only the functional features, but all the features of the noun head are copied in Spec DP, another way of saying that the noun head is overtly raised to Spec DP.²³ An overt chain is thus created, all links of which but the highest one must be deleted (see Solà 1996). Hence (6):

$$(6) \quad [_{DP} \text{ chat } [_{D^{\circ}} \text{ LA}] [_{NumP} \text{ chat } [_{Num^{\circ}} \emptyset] [_{NP} [_{N^{\circ}} \text{ chat}]]$$

Both in terms of economy of derivation and of isomorphism, it is clear that (6) deviates significantly from optimality. What causes the deviation is a particular morphological feature of Haitian, currently accounted for by assigning a strong D feature to the determiner head, which thus has the property of overtly attracting its complement to its checking domain (i.e., its specifier).²⁴ The straightforward spelling out of (5), on the other hand, would be optimal (with the usual provisos) since it only involves costless operations (merging and feature copy) which are driven by the sheer necessity of integrating the expression (see Collins 1997).

The structure schematized in (6) then feeds MF, with the special consequence that the phonological form of the determiner, so far noted as an archimorpheme LA, varies according to the last segment of what precedes it, hence the well-known alternations exemplified by *chat-la* ‘the cat’ vs. *pyebwa-a* ‘the tree’ vs. *moun-nan* ‘the person’ vs. *men-an* ‘the hand’. Notice that these alternations induce yet another distortion from isomorphism, since what precedes the determiner is not necessarily its whole complement in syntax as in *chat-la*, but it may be a subpart of this complement as in *chat ki souri* ‘the cat that smiled’ (see fn. 23), so there is no mapping from the syntactic relations (involving [chat ki souri] and LA) to the MF relations (involving [souri] and LA).

The plural correspondent of (5) is even more “imperfect”. Consider (7):

$$(7) \quad [_{DP} \text{ FF}_N [_{D^{\circ}} \text{ LA}] [_{NumP} \text{ FF}_N [_{Num^{\circ}} \text{ yo}] [_{NP} [_{N^{\circ}} \text{ chat}]]$$

with *yo* the plural morpheme paired with singular \emptyset . Again, straightforward spelling out yields an ungrammatical sequence **la yo chat*. What makes things even more complicated than in the singular, however, is that, depending on dialect, two grammatical surface forms correspond to (7), namely *chat-yo* and *chat-layo*, both meaning ‘the cats’. There is no problem deriving the latter from an intermediate structure similar to (6):

- (8) $[_{DP} \text{ chat } [_{D^\circ} \text{ LA}] [_{NumP} \text{ chat } [_{Num^\circ} \text{ yo}] [_{NP} [_{N^\circ} \text{ chat}]]$

But how do we derive *chat-yo*? The only way, it seems, is to assume optional merger of D° and Num° [+plural] at MF as represented below:

- (9) $D \cap Num[+plural]$
 $\swarrow \quad \searrow$
 $D \quad Num[+plural]$

where $\{D \cap Num[+plural]\}$ is the intersection of $\{D\}$ and $\{Num[+plural]\}$, whereas in the absence of merger no operation involves LA and *yo*, which are simply linearized.²⁵ Given the definition of merger (see Halle & Marantz 1993:116), $\{D \cap Num[+plural]\}$ has to be realized as one vocabulary item, i.e., in the present case, either *la* (or one of the other allomorphs), or *yo*, or a third item distinct from both. This latter option is the one exhibited by French where $\{D \cap Num[+plural]\}$ *les* [le] is distinct from D° /l/ as well as from Num° [+plural] /s/. It is not available in Haitian.²⁶ Neither is the first option, for the simple reason that it would make the $[\pm plural]$ contrast entirely invisible. Realizing $\{D \cap Num[+plural]\}$ as *yo* remains then as the only possibility. The result is a double departure from optimality. First, we have a vocabulary item that corresponds to two lexical items as well as to two functional categories, a multicategorical item in that sense. Secondly, this vocabulary item is ambiguous since it can mean either plural *tout court*, or plural associated with definiteness.²⁷

Other strategies for definiteness and number marking exist, but none convincingly closer to perfection. For instance, Martiniquais-Guadeloupean *se chat la* ‘the cats’ differs from Haitian mainly in the identity of the overtly raised phrase, namely the $NumP$ $[_{NumP} [_{Num^\circ} \text{ se}] [_{NP} [_{N^\circ} \text{ chat}] \dots]]$ instead of just the NP (see Bernabé 1983; Gadelii 1997). Jamaican *di pus dem* (we smile) ‘the cats (that smiled)’, on the other hand, implies overt raising-adjunction of N° (not NP) to Num° and covert raising of N° - Num° to D° (see Bailey 1966). In fact, Jamaican functions just like English in this respect, except for morphological details. In Ndyuka (see Huttar & Huttar 1994:202-203), we have the contrast *a pasi* ‘the path’ vs. *den kabiten* ‘the captains’ clearly indicating merger of Num with D as in Haitian, but with covert rather than overt NP movement. The same process accounts for Mauritian *ban lera* ‘the rats’ vs. *sa lera* ‘the rat’ (see Baker 1972). In the last two examples, imperfection is due entirely to the merger of the number morpheme with the determiner, which precludes one-to-one correspondence of syntax with MF.

I will limit myself to these few examples. They suffice to show that creole languages do not seem to be significantly more perfect than non-creole languages — a statement to be qualified later on. Meanwhile, it may be interesting to remark that we have reached this conclusion from an examination of the noun phrase, whereas studies conducted in the “simplicity” or “semantic transparency” paradigm have had a tendency to focus on the verb phrase. At least, this is my feeling, perhaps mistaken, and we shall have to see to what extent it does correspond to a significant observation.

As for question iii), Ndyuka and Mauritian have just provided us with two rather clear cases of suppletion. Indeed, *a* and *den* as well as *sa* and *ban* bear no phonological relation to each other although they belong to the same paradigm with *a* or *sa* [+def, -pl] and *den* or *ban* [+def, -pl]. This is genuine suppletion, since vocabulary insertion can only take place after the merged head Num^o-D^o has been formed, with entirely different items corresponding to different feature compositions of the merged heads. What does not seem to occur in creole languages, on the other hand, is suppletion of nonfunctional or not entirely functional morphemes such as English *go-went* or the Xârâcùù and Ancient Greek examples given earlier. Perhaps a more thorough inquiry than I was able to conduct for this article would uncover relevant cases, but it is doubtful.²⁸

The provisional conclusion we have reached, therefore, is that creole languages are not perfect languages. The next question is then why, and I will now try to provide an answer.

5. Why are creole languages not perfect languages?

As observed at the beginning of section 3, the above question makes sense only in reference to a particular creolization theory, the LBH. If we hold, on the contrary, that creole languages are relexifications of their first speakers' native languages, then there is no reason for them to be more perfect than the latter. The same is true if creole languages are viewed as more or less extreme cases of “ordinary” language change. Yet, there remains the nagging notion that creole languages are somehow “simpler”, i.e. more optimal, than their non-creole peers, which cannot be so easily waved away. To that extent, then, this question is a valid one calling for a nontrivial answer.

One answer, mentioned earlier, is that creole languages knew perfection

at their origin, as predicted by the theory, but then veered away from that state through ordinary processes of language change. There is naturally truth in this. However, as already suggested, the time depth does not seem to be sufficient to account for creoles' observed deviations from perfection, unless we make the additional assumption that these deviations were very sudden and early, with the implication that perfect languages are somehow not viable in natural environments. Let us keep this idea shelved a little longer.

The other answer that offers itself to me relates to a little-noticed mystery regarding the LBH. Recall the assumed scenario: children born in the degenerate linguistic environment of a pluriethnic labor force were not presented with coherent enough primary data to exert their Language Acquisition Device (LAD) on them.²⁹ Consequently, they resorted to their innate language bioprogram or C_{HL} to organize the data according to a grammar that, given the hypothesis, should present unmarked values for all or a significant majority of parameters. C_{HL} , however, did not provide them with a ready-made vocabulary. Precise terminology is important here, so I will insist on it. C_{HL} does include a lexicon, i.e. an innate set of conceptual and logical features bundled up as lexical items, that are then paired with audible vocabulary items or words that must be picked from the environment. Let me then ask the following question: how did it happen that the children never made mistakes in the pairing?

By this question, I mean the following: assume children are born with the feature bundle corresponding to the substantial concept *water'* and the logical feature <definiteness> or D, both lexical items.³⁰ The Jamaican word for *water'* is *waata*, a transparent variation on English *water*, which also expresses the substantial concept *water'*; similarly, the Jamaican definite determiner is *di*, again a transparent variation on English *the* or *this*, also definite determiners (the latter one with a deictic meaning attached). This seems self-evident, but it is not. If the linguistic environment was so degenerate, why didn't the children pick up any sound sequence they chanced to hear to clothe their innate lexical items? Or why didn't they make up their own vocabulary, something that has been observed in the private language of twins?³¹ More radically, why didn't children in such an untractable environment simply give up oral language for sign language, which seems to be just as innate to the human species (see, e.g., Pinker 1994)?

Those are conundrums the force of which is compounded when we realize to what extent the pre-creole linguistic environment was degenerate

according to the LBH scenario:

What happened in Hawaii was a jump from protolanguage to language in a single generation. [...] creole languages form an unusually direct expression of a species-specific biological characteristic, a capacity to recreate language in the absence of any specific model from which the properties of language could be “learned” in the ways we normally learn things (Bickerton 1990:171).

According to this quote, what strikes the children’s ears in a pre-creole situation is not language, but the productions of an evolutionarily more primitive system called “protolanguage”. I accept the reality of protolanguage as “a mode of linguistic expression that is quite separate from normal human language and that is shared by four classes of speakers: trained apes, children under two, adults who have been deprived of language in their early years, and speakers of pidgin” (Bickerton 1990:122). Obviously, protolanguage has something like a vocabulary. The question is: does it have a lexicon? Insofar as the lexicon is a computational component of language geared to LLF, the answer has to be a flat “no”. That is to say, the words of protolanguage should be considered signals, rather than signs in the Saussurean sense, directly expressing elements of conceptual structure in the broader sense.³² By implication, they should be sound chunks, unamenable to phonological and morphological analysis (see Bickerton 1990:144-5). In his 1998 article Bickerton further speculates that the crucial change (not necessarily a mutation in the strict sense) that turned protolanguage into language some 120,000 years ago in Southern Africa was the creation of a direct connection between two components already present in the protolanguage: thematic (theta) analysis and the phonetic representation of words (see his diagrams on pp. 347 and 352).³³ Theta analysis, probably an evolutive development from social intelligence, already enabled our pre-human ancestors to construe their environment in terms of situations and participants (“Who did what to whom?”). It can be seen as the bud from which syntax grew, insofar as it imposes a linear ordering and a hierarchy on the signal strings. In my terms, this connection is what permitted the emergence of a lexicon systematically linked with a vocabulary. Before the change occurred or when the connection is not activated, there is neither a lexicon nor a vocabulary (in the technical sense), but a stock of phonetic signals uniquely connected with the conceptual system and stringed according to principles extraneous to syntax. For clarity, I will call such a stock a “glossary”.

What does this imply for the question of the accuracy of the children’s

lexical choices when making the jump from protolanguage to language? As far as “denoting” items are concerned, perhaps not much. For instance, protolanguage must have had a glossary item for WATER, pronounced variably according to circumstances and resulting from the reduction of a vocabulary item of some language to the status of signal.³⁴ Insofar as they were able to recognize the denotation of this signal, children had no difficulty in transmuting it back into a lexicon and a vocabulary item of their new, creole language.³⁵ We only have to grant the fact that the shape of the signal was stable enough, which is open to debate given the absence of phonological structuring and the possible “macaronic” character of the environment.³⁶ But let us grant it. Difficulties really begin with nondenoting or “grammatical” items. The issue is a complex one, and we should be cautious not to oversimplify it.

By its very definition, protolanguage must lack such items altogether. What it may have, and apparently does have, are deictic signals translatable as “here”, “there”, etc., quantity signals translatable as “one”, “many”, etc. and other such things directly related to experienced reality. Given this, we have a possible scenario for the Haitian definite determiner /LA/ for example. All we have to assume is that the particular instantiation of protolanguage out of which Haitian emerged included a deictic signal pronounced « la » based on the French deictic adverbial *là* “there” and indicating remote location.³⁷ This signal would have been the seed, so to speak, both of the Haitian deictic adverbial *la* “there” and of the definite anaphoric determiner /LA/. The first of these evolutions seems straightforward enough, as it does not involve sweeping changes in the denotational content or the distribution of the item. The second one, in contrast, is not so obvious because it seems to be related to a property of the item that can only be true of French *là*, not of French-based protolanguage or pidgin « la ». I am referring to expressions such as *le gars là* ‘the guy (in question)’, where *là* postposed to a noun phrase ceases to be locational and is imparted with anaphoric force.³⁸ Such an expression would seem to be the direct ancestor of the Haitian equivalent *nèg-la* ‘the guy’. This raises several questions.

First, why would the creolizing children look for this item to create a definite anaphoric determiner? A possible answer is that an optimal natural language needs an overt definite anaphoric determiner. That is to say, C_{HL} requires that the universal lexicon item D should be mapped onto a vocabulary item, at least as the unmarked or optimal case. Although unprovable, this is a possible assumption, against which it would be irrelevant to rally the very

numerous languages, including creoles, that do without an overt D element, because there is no connection between optimality and frequency as a matter of principle. A stronger objection is that we already know Haitian to be less than optimal in this area, since the linear order of the determiner and the NP does not reproduce what the LCA entails. Why then did the creolizing children not go all the way if they were driven by the perfection requirement of C_{HL} , switching the order to produce the isomorphic **la nèg*? Recall moreover that linearization is supposed to be merely functional and in many cases irrelevant in the protolanguage (see Bickerton 1990:123), so that both orders « x la » and « la x » occur in it. It is therefore a bit surprising that the children would have kept the order attested in the source language instead of the reverse order they encountered in the protolanguage and which happens to be isomorphic once grammaticalized.³⁹

Another question that is raised is how did the creolizing children access precisely this item as a candidate form for D? Of course, there is a passable mental path from deixis to anaphora. The point is that it is not necessary, especially if the transition from protolanguage to language is assumed to be catastrophic rather than gradual, as it seems it must be. In other words, two alternative processes seem *a priori* equally probable. On the one hand, the children could have picked up any item they chanced to hear in the vicinity of what they recognized as a noun — and this means *any* item, given the syntactic unstructuredness of protolanguage. They might even have decided, upon hearing « neg la », that « neg » was to be the determiner and « la » the noun for “guy”. Lest this would seem far-fetched, recall that we are supposedly dealing with signals, i.e. sound chunks that do not have a meaning strictly speaking, but only an application to objects of the world. Once you switch to a neurally different system, there is no necessity for this application to be preserved.⁴⁰ Another possibility would have been for the children simply to invent words, using, say, *smurf* for a determiner. This takes us back to the twin language issue. In an environment such as the one we are assuming, creolizing children would probably form a tight small group, pretty much cut off from the adults. Imagine something like the situation described in Golding’s *Lord of the Flies*. Conditions for a private language would thus seem to be met. According to Bickerton (1990:191), twin languages are closer to protolanguage than to real language, although they “contain more grammatical items”. Consequently, there would be a two-step process: first the children realize their own protolanguage, possibly quite different from the one used by the

adults, then they make it into a language. The transition from one step to the other one would probably be very fast. Naturally, none of this ever happened: creole languages do not contain invented words or determiners related to words meaning “guy” or “hat”. It must be significant, however, both that it did not happen, as well as that it seems preposterous to even entertain the thought that it might have.

With the Haitian determiner, we saw a possible route from the source language to the creole through the protolanguage, with the problem that, given the hypothesis as it stands, the implied process lacks necessity, which reduces the account to being a description of the facts rather than an explanation. In other cases, the process itself is harder to imagine. Indeed, given Bickerton’s (1989:176) dictum that “the syntactic properties of creoles result from the interaction between universal principles of syntax and the (often highly impoverished) properties of the vocabularies that creole languages inherited”, we have to ask ourselves: what happened when the creoles did not inherit anything in a particular domain?⁴¹

Consider, for instance, the category of complementizers. Since protolanguage ignores embedding, no such items are expected to exist in it (see Bickerton 1990:125-6). In contrast, complementizers are clearly a requisite of C_{HL} , i.e. necessary items of the universal lexicon, which needed to be expressed. The question then is how is it possible that, in so many cases, the creolizing children hit upon a form that is already a complementizer in the source language or is at least compatible with that function? In a number of instances, the answer may be that the item had actually survived in the pidgin, stripped of its function, but used in contexts which allowed for its reinterpretation. Such may be the case of Haitian *lè* ‘when’ from French *l’heure* ‘the hour/time’. It is quite conceivable that the French-based protolanguage had a deictic signal « *le* » somehow used to set the situation in time and derived from some complex embedding expression more contentful than plain *quand*, probably *à l’heure que* ‘at the time that’. But what about more abstract items such as *ki* from French *que* functioning as a relative pronoun in Haitian and also as a complementizer in Seselwa? Obviously, we must assume it was present “in disguise” in the antecedent protolanguages. But then why and how did the early creole speakers precisely zoom in on this signal or portion of a signal to spell out the lexical item C if they had no access to the source language, it being the only system where the sound sequence /kV/ can be identified as a complementizer?⁴²

Tense-aspect markers are also relevant to this discussion. In some cases, they conform neatly to what the LBH predicts. For instance, Tok Pisin *bai* meaning [Future], from an earlier *baimbai* from English *by and by*, clearly descends from a protolanguage deictic word-signal « *baimbai* » used to locate the situation in the not-yet section of chronological time.⁴³ Given this, parsimony would lead us to expect Seselwa *(a)pe* to be a Future marker because its etymon *après* “after(ward)” is probably the best French candidate to serve as a “later” deictic signal.⁴⁴ But *(a)pe*, like its Haitian equivalent *ap*, is a Progressive marker in direct continuation of the French dialectal expression *être après faire* ‘to be doing’ (see Corne 1977), whereas Future in Seselwa is marked with the preverbal auxiliaries *a(va)* or *pu*, both continuing equivalent French forms — compare *Elle va partir* ‘She’s going to leave’, *Elle est pour partir* ‘She’s about to leave’, the latter dialectal. It is highly implausible that such forms were part of the glossary of the antecedent protolanguage. The same argument applies to the Neo-Caledonian French-based creole Tayo where the Future marker is *va* and the Progressive marker is *atrade*, a direct inheritance from French (*être*) *en train de* (see Ehrhart 1993; Kihm 1995).⁴⁵

On the other hand, the Seselwa and Tayo completive markers *(fi)n* and *(fi)ni*, both from French *fini* ‘finished’, correspond well with what one expects to find in a pidgin glossary (compare Tok Pisin *pinis*) and have no direct correspondent in French.⁴⁶ But take the Past marker */ba/* of Guinea-Bissau Kriyol as in *N bay ba* ‘I had left’. To assign it an origin in the protolanguage, we have to assume that */ba/* represents the final syllable of an item */kaba/* from Portugues *acabar* ‘to finish’, so that the pidgin equivalent would be something like « *mi bay kaba* ». The problem with this supposition is that *kaba* is a Kriyol word meaning either ‘to finish’ or ‘and then’, that is never used to express accomplishment or past, and that */ba/* finds a much more straightforward etymon in Portuguese */-va/*, viz. the 1st and 3rd person singular indicative past morpheme for first group verbs.⁴⁷

What I want to show with these examples is that, universal factors being granted, the properties of the items we can reasonably suppose to have been present in the protolanguage do not suffice to explain the creole outcome. Properties of the source language at least, and probably of other languages as well, that never could go into the protolanguage as defined, must also be taken into account. We cannot explain otherwise the quality of the fit between the source and creole vocabularies. Naturally, this is a problem for the theory.

Another, more radical, problem for the theory has to do with economy of

derivation. This could have been raised at other points in this study, such as section 2, but I think it especially appropriate to raise it here in view of what must follow. Economy of derivation means that, whenever several derivations are possible from one numeration, none of them leading to a crash at PF or LF, the most economical one is selected over its competitors which are thus blocked and rendered unacceptable (e.g., Kitahara 1997). Assessing which derivation is the most economical cannot be done in an absolute fashion, however. Insofar as the lexical items that compose a numeration have morphological properties, these properties must be entered into the computation, which is why “most economical” (in the sense of, e.g., “involving the least number of steps”) is always a relative evaluation, and why natural languages are not optimal. Let us place ourselves, now, in the mind of a child who is starting the process of producing a real language from UG using material from the protolanguage she is hearing. Since this material cannot be assumed to have properties relevant to UG, it follows that in this case, and in this case only, “most economical” should be understood absolutely. That is to say, derivations could involve only “no cost” steps, and costly derivations such as the one leading to Haitian DP would never have a chance to arise — at this stage.

The foregoing argument may be recast even more perspicuously, perhaps, if we adopt the Optimality Theory (OT) version of Minimalism (see, e.g., Grimshaw 1997).⁴⁸ In this framework, particular derivations are driven by the relative ranking of more or less violable universal constraints. Constraint ranking is language-specific, and it is what the child discovers in the primary data she is presented with. In a situation where there is no language to proceed from, therefore, constraint ranking would have to be absolute. As far as I know, researchers of OT have not envisaged such a case yet. A consequence one can imagine is that all constraints would then rank equally, being either all violable, which would lead to chaos, or all inviolable, which would result in total blocking. Presumably, one would have to supplement the theory with a measure of markedness whereby one ranking is considered unmarked with respect to all others. I don’t know whether this is feasible, but it is safe to surmise that such an unmarked ranking would yield isomorphism.

Should we take these problems as indicative that the question that heads this section could simply be answered along the following lines: creole languages are not perfect because the situation that could have made them so never occurred? That is to say, there was no chaos;⁴⁹ the children heard (also)

language, not (only) protolanguage, and they were presented with primary data that were certainly degenerate as they always are to some extent, but that were also structured enough that they could deflect the children's innate grammar according to specific parameters. These data would have come both from the parents' first languages, i.e. the substrate, and from the more or less pidginized superstrate they had acquired or were acquiring. From there, one might proceed to the conclusion that the LBH is actually falsified, since the scenario upon which it crucially rests can be shown not to be realistic. The possibility would then be opened that it was not the children who played the central role in creolization, but the adults, such that a different process, e.g. relexification, was the leading force.

Actually, such an answer may be a little too simple. What it fails to account for is the other unavoidable observation about creole languages, namely that, although they are not perfect, they may be closer to perfection than the common stock. In the following section, I present some highly speculative reflections on this topic.

6. Why are creole languages less imperfect (if they are)?

Recall first what would make a language perfect, given a substantive and hopefully falsifiable theory of human language such as Minimalism. Since *l'arbitraire du signe* is our fate, a real (as opposed to "ideal") perfect language could only be one with no more than two components: a narrow syntax or LLF and an MF-PF component where each element of LLF would be sounded in the exact location assigned it by LLF. In other words, LLF and MF-PF would be isomorphic.

This is precisely what the common observation that creole languages have little morphology means: LLF and MF-PF tend to be isomorphic.⁵⁰ As I tried to show in section 4, this observation is not entirely true. Yet, it holds to a large extent. Examples aiming to demonstrate this point are superfluous (see DeGraff 1997 for a lucid exposition of the issue in relation with the verb phrase). A question worth discussing, in contrast, is what we mean by a tendency to isomorphism. After all, a mapping is isomorphic or it isn't. What we need, therefore, is a measure of the degree to which some mapping or a given set of mappings deviate from isomorphism, i.e. perfection. Two considerations make such a measure possible.

First, there is the fact that language is modular to a high degree. This is true not only in the usual sense that language consists in several articulated systems, but also in the sense that different syntactic categories may behave differently. For instance, we saw that it is mainly the noun phrase that accounts for deviations from isomorphism in Haitian. The verb phrase, in contrast, is isomorphic due to the absence of overt verb movement explaining why, for instance, the negation *pa* precedes all verbs in Haitian, while *pas* follows finite verbs in French (see DeGraff 1997). Consequently, a language where only some categories deviate may be said to be globally more isomorphic than a language where all categories deviate.

Secondly, morphological processes responsible for dysmorphisms fall into two classes: fusion and merger: “Merger, like head-to-head movement, joins terminal nodes under a category node of a head (“a zero-level category node”) but maintains two independent terminal nodes under this category node [...] On the other hand, fusion takes two terminal nodes that are sisters under a single category node and fuses them into a single terminal node [...] Unlike merger, fusion reduces the number of independent morphemes in a tree” (Halle & Marantz 1993:116).⁵¹ To the extent that morphemes correspond to functional categories, fusion is clearly more dysmorphic than merger. Now, fusion phenomena seem pretty rare in creole languages. Haitian *yo* in the dialect where it spells out both number and definiteness is an example. There are others, but not many. Merger, on the other hand, when it occurs, is mainly of the most “superficial” type in which the order of the constituents is preserved — see Haitian *tap* merging Past *te* and Progressive *ap*.⁵² Guinea-Bissau Kriyol causative verbs as in *firbinti yagu* ‘to boil the water’ (vs. *yagu na firbi* ‘the water is boiling’) is a case where merger entails a disruption of the underlying order where the functional node corresponding to the causative morpheme presumably dominates the verb node (see Kihm 1994; on causativization, see Baker 1988). But note that the individuality of the morphemes is well preserved, unlike fusion cases such as English causative *boil* or *fell*. Moreover, such morphological processes are extremely limited in the creole family. In particular, they seem never to occur in plantation creoles.⁵³ Both features — rarity of fusion processes and limitation of merger processes — thus concur to make creole languages as a whole more isomorphic than languages where these processes are extensive.

An objection one could raise (one I have myself raised) is that these features also characterize a number of apparently non-creole languages. To it,

I can now see two answers. First, the more I look into linguistic diversity, the more I am convinced that creole languages rank distinctively high on the isomorphism scale, despite occasional departures. Indications pointing in that direction have been given above.⁵⁴ Secondly, the fact that a language group we know to be distinguished by the special way it came into being is characterized by these features must be significant, even though a few languages not in the set share them to some extent.⁵⁵

All this points towards the conclusion that the notion of relative distance from perfection is amenable to computation.⁵⁶ We may therefore be able to substantiate the apparently true observation that creole grammars are consistently less remote from perfection than many (perhaps all) other grammars are. Of course, we may discover that the observation is false or even meaningless. In the framework of the current theories of grammar and given our present knowledge of the relevant languages, however, such an outcome seems improbable.

If the observation is true, then, we are facing a dilemma. On the one hand, we have a theory, the LBH, that predicts that creole grammars should be perfect, and that does not explain why they are not entirely so. On the other hand, we have a second theory, substratism, which explains why creole grammars are not perfect, but does not predict that they are less imperfect than many (perhaps all) other grammars.⁵⁷ Clearly there is a missing link. My suggestion is that the missing link might be found in an idea that was mentioned twice in passing earlier, namely that perfect grammars may not be viable.

Let us assume that the LBH is true to the extent that, in typical pre-creole situations, transmission from the parents to the first generation children was not what it ought to be. Several reasons concur to it having been the case: linguistic (the parents' protolanguage-pidgin was not an adequate input), psychological (the parents were in a state of shock, to put it mildly), sociological (the parents were simply not able to take care of the children), and perhaps other. The consequence was that the children's C_{HL} had to run in neutral, so to speak, i.e. to function on its own inner resources to an extent that it never has to in ordinary situations. Had it done so entirely and long enough, the result would have been a perfect language with a randomly selected or invented vocabulary. Why didn't this happen? The primary reason — at least this is my proposal — is that perfection is a virtual property of language as a computational system for representations. But, as repeatedly mentioned, that system is

interfaced with other systems which Chomsky (1993) calls “performance systems” having to do with articulation-perception, general cognition, intentions, etc. For these systems, which take care of the communicative aspect of language, roughly speaking, perfection is not only irrelevant, it would probably be detrimental. The gist of Chomsky’s quotation at the beginning of this paper is therefore that these extraneous systems force the computational part of language to deviate from virtual perfection, and that it cannot be otherwise.

Given this, the other half of the argument proceeds from the fact that, as we saw, there is a logical flaw in the claim that “the syntactic properties of creoles result from the interaction between universal principles of syntax and the (often highly impoverished) properties of the vocabularies that creole languages inherited” (Bickerton 1989:176). Indeed, taking the protolanguage hypothesis seriously, these vocabularies (or glossaries as I prefer to call them) simply did not have properties relevant to real language. This leaves only the possibility that real language was also present in the situation. Of course, this is a possibility that Bickerton freely admits on empirical grounds (see, e.g., Bickerton 1992). What I wished to demonstrate is that it is in addition a logical necessity that it was.

Under what form was real language present? The easy answer — being easy does not prevent it from being true in many cases — is that the original linguistic communities did not vanish entirely “over there”, even though they broke down to a large extent.⁵⁸ But let us grant the extreme case, that nothing survived of these communities, and that only protolanguage-pidgin could be resorted to. Even so, there were niches for real language. One is the mother’s mother tongue, which she may have talked to her baby in while caring for it, even if to nobody else, a possibility generally forgotten. Recent studies strongly suggest that such early imprints, even before the language faculty begins to manifest itself, perhaps even before birth, are crucial for the child’s later linguistic development (see Pinker 1994, in particular p. 264). Psychologists working in this area seem to have only considered the phonological dimension of language so far. Would it be delirious to speculate that other dimensions of UG might be affected as well, implying that children do not merely perceive their mothers’ speech as language sounds, but as language signs as well? I think not, although this is naturally for specialists to determine, provided experiments can be devised. Imagine then a number of mothers speaking the same or typologically similar languages to their babies; even if they were scattered over an area and not in communication with each other

(or only through the pidgin), a “crypto-substrate” would ensue that might show once the children begin to address each other in the emerging creole. In this way, protolanguage signals could be transmuted into language signs along specific paths, for instance with specific lexical structures. I repeat, this is only an extreme assumption, just for the case we prefer to reject the possibility of early substratal or superstratal influences, for empirical and/or theoretical reasons.

At the level of abstraction at which I am comfortably sitting, it does not make much difference where the real language data actually came from. What is important is that the children’s language faculty did not run in neutral for long, but there was something to deflect the children’s innate grammar from perfection in a certain direction, given that it had to deflect — or die — in any event. On the other hand, the fact that the children’s innate grammar passed through this neutral running phase continued to show in the greater degree of LLF-MF-PF isomorphism of the resulting creole.

But is not such a phase an integral part of all linguistic acquisition? Yes, indeed. In ordinary acquisition, however, it passes off without leaving a special trace, since children end up with a full knowledge of the language their C_{HL} has been given to process. What makes creole situations different, therefore, is that children do not end up with a full knowledge of the language(s) the primary data they have access to come from. After all, with due apology for the truism, Haitian is neither Fongbe nor French, even though both played a role in its formation, it is Haitian, a specific parameterization of UG. This is not yet enough, however. Indeed, children did not end up knowing a mixed language either, something not closer to perfection than any of its composing parts. What is really different, therefore, is that the way the primary data were acquired just as real language material (not as portions to be completed of full languages) left room for the neutral running phase to leave its imprint on the final system. There is nothing necessary in this. In fact, as with any natural phenomenon, it needed a very subtle balance of the factors involved to yield precisely this result, a creole, rather than one of the other possible outcomes: full knowledge of the substrate language(s), full knowledge of the superstrate language, mixed language, extended twin language, sign language, total muteness, to list them (perhaps not exhaustively) in decreasing order of plausibility.

Needless to say, I am quite unable to assess this balance, the factors entering into which varied from situation to situation, perhaps giving more prominence to universal factors here, to substrate factors there, etc. I am

merely suggesting that, between all-universal and all-substrate theories, there is room for what I would like to call a Relativized Universal Grammar Hypothesis (RUGH) that may be a bit more adequate in terms of acknowledging the complexity of real life. In the RUGH as in the LBH, unparameterized UG played a crucial role in shaping the syntactic form of creole languages. Yet, unlike the LBH, the RUGH also takes into account logical and empirical arguments that show that UG cannot have remained unparameterized for any measurable amount of time. This means that the original environment was not and could not have been pure protolanguage-pidgin, but contained real language as well, something the LBH acknowledges as a fact, without drawing the crucial consequences that it entails.

The difference with non-creole situations resides then in the way the child’s language faculty reacts to the environment it finds itself in. I am speculating here (recklessly perhaps) that the language faculty includes a capacity to sense whether the linguistic environment is “normal” or not. That is to say, a child should be able to recognize whether she is hearing only real words that match effortlessly and automatically with her innate lexicon; or whether she is given to process a mixture of real words and non-words, i.e. protolanguage signals. Of course, it is for psycholinguists to tell us whether such an assumption is plausible, unlikely, or inane. It seems to me, however, not to be quite out of touch with what we think we know about the human brain. Imagine, for instance, a child who grew in an environment with an unusually high proportion of slant lines. We do not know for sure what would happen — and we do not want to know — but it seems reasonable to suppose that her visual cortex would have a lot of extra work to do in order to mature normally. In the same way, C_{HL} in a creole situation has extra work to do because there are too many things out there that do not fit its genetic specifications.

Given this, a necessary condition is that there is more than one child. I agree with Bickerton (1992) that quantity *per se* is not at issue. But it seems to me obvious that there should exist a community, however small, of children sharing the same predicament if there is to be an incentive for rapidly creating a real language. An isolated child or (more realistically) scattered children unable to regroup in the abnormal situation we are assuming would a) be seriously impaired in their linguistic development; b) end up speaking one of the real languages fragmentarily present in the environment.⁵⁹

Finally, I reject the notion that there could be only protolanguage present

in such contexts. I reject it for empirical reasons because C_{HL} would then have to run fully by itself, which would give rise to a perfect language that could not evolve into the imperfect creoles that we find. I also reject it on theoretical grounds, because protolanguage signals would be intrinsically unfit to match the children's innate lexicon, perhaps forcing them to switch to an alternative device to realize this lexicon, such as signing. In other words, they would be in the situation of deaf children who spontaneously evolve sign language because their handicap denies them access to oral language. There must be at least some real oral language available to trigger the children's "choice" of this means to express themselves.

7. Conclusion

It may be true, then, that "the universalist and substrate hypotheses complement each other", as Mufwene (1986) puts it. This is not to be taken in an "anything goes" or "a bit of everything" spirit.⁶⁰ If they do complement each other, it is in the sense that their respective domains of validity may correspond to distinct layers in a more fine-grained stratification of creolization processes than has been presented so far. Whether such a fine-grained stratification can really be substantiated — whether in particular we will ever be able to evaluate the precise balance of the factors involved in different concrete situations — is of course a different matter. Its reality, however, is what may allow us to provide an answer to the leading question of this article: creoles are not perfect languages, but they live on with the marks of perfection that the unnatural conditions of their birth forced on them.

Assuming this, two main tasks follow. The first one is to ascertain what degree of validity the hypothesis may have, and it is not for (traditional) linguists only to do it. Studies of acquisition and sign language, in particular, have a crucial role to play. The essential separation of protolanguage from language is also something that can be contested. In line with current arguments about the crucial transformation that language represents, I have supposed it to be radical. Should it turn out not to be so, things would take on a different outlook. The presence of real language in the environment would become less of a necessity. On the other hand, protolanguage would then be demoted to the status of "paralanguage" (perhaps some sort of "foreigner's talk" or "broken language" as studied in Ferguson & DeBose 1977) necessar-

ily influenced by its users’ first language(s), which could therefore act as a substrate. As a result, the universalist hypothesis would break down or lose much of its relevance. What I have tried here is merely to push this hypothesis to its logical consequences, while indicating ways that my interpretation of it can be falsified. Therefore, to the extent that I have succeeded, and should my hypothesis prove untenable, then I think we should renounce assigning any special role to universal grammar in creole genesis.

The second task, supposing the hypothesis to be valid or at least plausible, is to define its coverage. Indeed, “creole” (or “pidgin” for that matter — see Hancock 1996) is as much a putative natural class as it is a mere label uncritically inherited from scholarly tradition. This means there is no *a priori* reason why all the languages we currently call “creoles” should fall under the RUGH or some variant of it. I am alluding, for instance, to the distinction between plantation and fort creoles for which it is not certain whether a) it is discriminating enough or b) it is one of kind or of degree. If of kind, then it is just possible that fort creoles are of no concern to creole studies strictly defined — i.e. studies of language growth in unnatural situations; if of degree, as I rather believe, then it takes us back to the issue of precisely assessing the balance of the factors involved, as far as it can be retrieved. Beyond polemics — surely a useful exercise as it forces participants to sharpen their arguments and makes conferences more lively, but one for which I have little taste — there is thus much room for discussion.

Notes

1. Another issue from which language perfection as understood here must be distinguished is that of philosophical or artificial languages in the Leibnizian tradition which aim to be perfect means for expressing the thought (see Laycock & Mülhäusler 1990). The necessary distinction rests on two things. Firstly, such “perfect” idioms proceed from more or less explicit theories of the global organization of the mind which their proponents assume to be true. In the present perspective, in contrast, perfection is referred to a theory of the language faculty — i.e., an autonomous component of the mind — which is empirically founded on the observation and analysis of natural languages, and is thus permanently open to falsification. Secondly, a point that brings us back to the difference between perfection and adaptation, philosophical languages have proved entirely unsuitable for human communication (and artificial languages such as Volapük or Esperanto as well, unless they are heavily “naturalized”). Our view of perfection, on the other hand, entails that there *cannot* be perfect natural languages in the real world. The question, therefore, is whether imperfection is absolute — rendering the notion useless — or

relative. I am grateful to Peter Mülhäusler for remarks that led me to this tentative clarification.

2. In the sense that these systems should be able to make something (sense, sound) from what language delivers to them.
3. I use “optimal” as synonymous with “as perfect as possible”, therefore simple from a computational perspective. It does not commit me to any version of Optimality Theory. Neither does it imply anything as to the (essentially non-linguistic) question of whether an internally optimal (simple) language is also optimal from the viewpoint of its uses in social and natural contexts.
4. In saying this, I am not addressing issues of cultural associations, what counts as a prototypical tree, etc., but keeping as firmly as possible to the level of denotation. Of course, I am aware this is a complex problem that deserves much more than a brief footnote (for a view of the extent of the problem, see Foley 1997). The important thing to keep in mind is that TREE is not English, or French, or Wolof, or whatever, but a symbol for an assumedly universal feature bundle, the determination of which is an empirical issue. In this way, languages where the word meaning TREE is a verb, if such exist, do not pose a problem for the framework, because syntactic categories do not have to be features of the lexicon, and probably are not or are underspecified.
5. We thus capture Benveniste’s valuable remark that the sign is arbitrary only with regard to the system, but not in relation to the native speaker’s psychological state (see Benveniste 1966, Chapter 4). I ignore, probably not harmlessly, the usual problems with language names (what is English?).
6. I am assuming that complex roots can be assembled out of simplex roots using the elementary operation Fusion.
7. More precisely, I assume the category (here “verb”) to be determined at the stage where the root moves to the “light” verb *v* in the syntax (see Marantz 1999).
8. I keep here to a conservative view of agreement involving an Agr projection. Alternative views as in, e.g., Chomsky (1995) or (1998) do not affect the argument.
9. See Dowty et al. (1981, Chapter 7) for the theory of types involved. Also see Chierchia (1985).
10. As is well-known, suppletion may be due to the bringing together of different roots into one paradigm, as in the English or Greek cases, or to sound changes that obscure the relation between alternant forms of the same root, as in French *oeil* vs. *yeux* “eye(s)”.
11. This harks back, of course, to the Neogrammarian conception of the relationship between sound change and grammatical change.
12. Sign Language studies should prove decisive in this area as in so many others.
13. Note that I am not taking sides here, even though I do not intend to maintain such a cozy aloofness to the end.
14. If the distorting factors were integral to C_{HL} , we would expect all languages to deviate from isomorphism to the same extent and in the same way, or to diverge wildly. This is more or less the view of those who think there is no limit to the ways languages may differ

from each other. The actual picture, however, seems more in line with the existence of a common core modified under the influence of diverse external factors.

15. Except insofar as they induce linguistic diversity which may have a positive value from the point of view of evolution, as has been argued for the different “dialects” of some bird species. But such an effect is only indirect and accidental. Linguistic diversity could still be attained in the absence of complex morphophonological processes, which retain their basic character as unmotivated (although not random) departures from perfection in the “formal” meaning of the term.
16. It is important that the reader should not think of this passage as an attempt to refute the bioprogram. What I am doing here is trying to make fully explicit the meaning of a notion which deserves careful consideration, whatever one’s final assessment of it.
17. Another name for the language bioprogram, not much used now, is “natural semantax” (see Bickerton 1974). It is interesting that the *a priori* more accurate term “morphosemantax” was never proposed.
18. I take the examples from Pustejovsky (1995:91).
19. Measured by this yardstick, French with the three words *montre*, *horloge*, and *pendule*, where English has only *watch* and *clock*, and Spanish only *reloj*, stands well ahead in terms of transparency.
20. See below for the reason why LA is capitalized.
21. That is to say, [_{Num}Ø] is a lexical item that has no correspondent in the vocabulary.
22. More precisely, movement conceived as the combination of the two elementary operations Copy and Delete (see Collins 1997).
23. Actually the head and its possible complements, i.e. the whole NP, as shown by, e.g., *chat ki souri a* ‘the cat that smiled’. In fact, given the restrictions on head movement, N° alone could only adjoin to D°. I assume movement, be it overt or covert, to be strictly successive cyclic, so that no intervening head may be skipped. The instruction “delete all links but the highest one” may be seen as a requirement of the syntax-MF intermapping, according to which arguments may occur only once.
24. Of course, “strong feature” is only a description with little explanatory value as long as it is not given a substantial content. This does not affect the logic of the explanation, however.
25. Since merger is optional in this case, we are free to assume it to occur or not to occur in the singular as well. Either way, the phonological effect is nil. Also note that D and Num[+plural] may be viewed as naming the sets of all definite entities and all plural entities respectively, hence the appropriateness of the Boolean notation.
26. Unless *layo* is considered a merged item distinct from *la* and *yo*. I do not see, however, how such an item could be distinguished from the mere linearization of *la* and *yo*.
27. Not taking into account its third meaning as third person plural pronoun, perhaps a separate homophonous item in present-day Haitian. Of course, if the *yo* and *layo* dialects are strictly separated, the ambiguity issue could never arise. I would be surprised, however, if that was the case.

28. Whether suppletion is considered to be inflectional or derivational, its opacity and semantic irregularity run afoul of all reasonable criteria for creoleness (see also McWhorter 1998).
29. Or their Language Faculty *tout court*, if there is no such thing as a separate LAD (see Chomsky 1998).
30. *Water'* is the translation into Intensional Logic of the English word *water*, and it may be considered the sense of this word (see Dowty et al. 1981:189ff.). The idea behind all this is that children know what water is before they can name it (even before they encounter it outside their mother's womb), a perfectly reasonable conclusion that follows from the model of the mind that underlies our present theory of language. Whether this innate concept will be later modified to some extent — probably never drastically — by the culture into which the child happens to be born is a distinct issue, as already mentioned.
31. In fact, twin languages include both words from the surrounding language and private words (see below).
32. By this, I mean the cognition of the outer world and of inner states that we share to some or a large extent with all types of humans, the apes, and probably other non-humans as well.
33. The new neural configuration so created did not erase the previous one, as proved by the fact that protolanguage is efficient to this day.
34. I use small capitals for the concept of the object, as distinct from the linguistic sense of the concept.
35. I take denotation in the Fregean sense of true or false when applied to concept *x*.
36. For instance, we know of situations such as that of the Sino-Russian Pidgin of Manchuria where more than one glossary was used concurrently, so that the signal for WATER was either something like [vada], from Russian, or something like [shuo], from Mandarin Chinese (see Lyovin 1969; Hancock 1996). Apparently, no children were born in that situation, but what if they had been? The issue is also linked with whether the crucial interaction in such settings was with the mother, which would probably warrant a higher degree of stability, or with the peer group. Given the very fact that creoles did appear, the latter would seem more plausible.
37. I use double angle brackets to indicate protolanguage words.
38. Because such expressions belong to colloquial registers, they are better understood in context, as in *Alors, le gars là, il m'a dit...* "Then the guy told me...", not synonymous with *Alors, ce gars-là, il m'a dit...* "Then that guy told me...". Anaphoric *là* is especially frequent in Quebec French.
39. Bickerton (1990:123) allows for the possibility of the word order of the source language to be transferred into the pidgin. But given the separateness of protolanguage from language in terms of organization, such a transfer can only be a matter of pure chance. What I mean is that it cannot be constrained, let alone expected.
40. Here, I am taking the hypothesis seriously and pushing it to its logical consequences. If these consequences look uncanny, then it may mean that the complete identification of protolanguage with pidgin goes too far. I will return to this.

41. Recall that, strictly speaking, the creole did not inherit a vocabulary in the linguistic sense of the term, but a glossary of signals. Neither is it very rigorous to claim that the *creoles* inherited: the children did, and we don’t really know what they heard. It is interesting that Bickerton seems to forget the intervening protolanguage stage in this quote.
42. It is indeed very possible that « ki » was present in the protolanguage only as part of a longer signal where it played no specific role; see, for instance, Bissau Kriyol *tok* “until”, a reduction of *torok* (also attested in more conservative varieties), itself ultimately from Portuguese *até a hora que* “until the hour that” (see Kihm 1994). It seems safe to assume that the [k] buried in the pidgin signal was no more than a final consonant, since it is still that in the creole word. The Haitian interrogatives *kimoun* “who?”, etc. are a different case, as it is not clear that they date from the pidgin stage.
43. Similarly Papiamentu *lo* from « logu » from Portuguese *logo* ‘soon’.
44. See colloquial French *Maintenant, j’écris; après, je me repose* “Now I’m writing, then I’ll rest”. French schoolteachers have been insisting for generations that *après* ought not to be used in this way, and that only *ensuite* is “correct”, to no avail fortunately.
45. Since *en train de* is a phonological unit, a quasi-word, in French, there is no intrinsic necessity to invoke pidginization to explain the Tayo form.
46. For instance, Seselwa *Mô n vini isi* “I have come here” (Corne 1977:107) is entirely different from *?J’ai fini de venir ici* which, to the extent that it means something, is interpreted as “I’m finished with coming here” as well as from *J’ai fini par venir ici* “I’ve finally come here”, the only verbal constructions involving *finir* in French.
47. The “and then” meaning of *kaba* is easily understood when one considers discourses such as *N tarbaja, kaba n diskansa* “I worked and then (having finished) I rested”.
48. I am not personally an adept of OT, but I think it is an interesting theory that is directly relevant to creolization issues. That all this discussion is highly theory-internal should deter no one, I think. Without theories, there would be little to talk about, in any event. We would certainly be closer to spiritual perfection, but that is a different issue.
49. Taking due account of Bickerton’s (1994:66) correction regarding this term.
50. The other common observation that all creoles have similar grammars follows from this proposition, since LLF is universal by definition.
51. In a more recent version of the theory (see Marantz 1999), fusion is reallocated to the lexicon, as being the operation that joins features into bundles, which are then delivered to MF, where merger may additionally occur. I keep to the older version, as this theoretical readjustment does not affect the argument.
52. Under the common assumption that Tense dominates Aspect. The implication that *tap* is a single Vocabulary item seems a reasonable one, given the absence of a purely phonological process systematically reducing such vowel sequences (see *li te atè* “s/he was on the ground” — not */li tatè/).
53. On the significance of the “plantation vs. fort” distinction, see Kihm 1995.
54. Also see McWhorter (1998). McWhorter points out three criteria, the *conjunction* of which sets creoles apart from non-creole languages, viz. lack or paucity of inflectional

processes (including affixation, apophony, and suppletion), lack of lexical or grammatical tone, and semantic regularity of derivational processes (when any). Here, I have only considered the first criterion, but the other two are clearly associated with it in the languages I took as examples.

55. The “special way” is an independent historical fact, so there is no circularity in the argument. The claim that all languages, or many more than we suspect, may have been creoles at some stage is bound to remain empty, unless it can be proved that the circumstances attendant on the genesis of the creoles that we know were reproduced at other time(s) in (pre)history, which strikes me as both unprovable and unlikely.
56. An interesting proposal is Sadock’s (1991) General Intermodular Homomorphism Constraint which formalizes such a measure. Although Sadock’s theory of Autolexical Syntax is seriously at variance with Minimalism or Distributed Morphology on many points, its axiomatic foundations are, I think, compatible with those of these theories (see Jackendoff 1997).
57. Theories that emphasize the role of language-change factors such as Thomason & Kaufmann (1988), or the role of the lexifier language such as Chaudenson (1992), align with substratism in this respect. Being variously necessary components of these theories, universals of second language (L2) acquisition take no autonomous part in the present discussion.
58. I am referring here to the situation of plantation creoles. With fort creoles, there was no “over there” to begin with, except perhaps in a sociolinguistic sense, because of change of religion or the like (see, e.g., Clements 1996).
59. The superstrate language is the surest bet in this case, since it is the one most likely to endure.
60. I am not suggesting that Mufwene took it in this spirit.

References

- Bailey, B.L. 1966. *Jamaican Creole syntax: a transformational approach*. Cambridge: Cambridge University Press.
- Baker, M.C. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.
- Baker, P. 1972. *Kreol: a description of Mauritian Creole*. London: C. Hurst & Company.
- Benveniste, E. 1966. *Problèmes de linguistique générale*. Paris: Gallimard.
- Bernabé, J. 1983. *Fondal-natal: Grammaire basilectale approchée des créoles guadeloupéen et martiniquais*. Paris: L’Harmattan.
- Bickerton, D. 1974. “Creolization, linguistic universals, natural semantax, and the brain”. *University of Hawaii Working Papers in Linguistics* 6:124-41.
- Bickerton, D. 1984. “The language bioprogram hypothesis”. *Behavioral and Brain Sciences* 7:173-221.
- Bickerton, D. 1989. “Seselwa serialization and its significance”. *Journal of Pidgin and Creole Languages* 4:155-83.

- Bickerton, D. 1990. *Language and species*. Chicago: The University of Chicago Press.
- Bickerton, D. 1992. "The sociohistorical matrix of creolization". *Journal of Pidgin and Creole Languages* 7:307-18.
- Bickerton, D. 1994. "The origins of Saramaccan syntax: a reply to John McWhorter's 'Substratal influence in Saramaccan serial verb constructions'". *Journal of Pidgin and Creole Languages* 9:65-77.
- Bickerton, D. 1998. "Catastrophic evolution: the case for a single step from protolanguage to full human language". In *Approaches to the evolution of language*, J.R. Hurford, M. Studdert-Kennedy and C. Knight (eds), 341-358. Cambridge: Cambridge University Press.
- Brody, M. 1995. *Lexico-logical form: a radically minimalist theory*. Cambridge, Mass.: MIT Press.
- Chaudenson, R. 1992. *Des îles, des hommes, des langues*. Paris: L'Harmattan.
- Chomsky, N. 1986. *Knowledge of language*. New York: Praeger.
- Chomsky, N. 1993. "A minimalist program for linguistic theory". In Hale and Keyser (eds), 1-52.
- Chomsky, N. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1998. "Minimalist inquiries: the framework". (MIT Occasional Papers in Linguistics 15)
- Chierchia, G. 1985. "Formal semantics and the grammar of predication". *Linguistic Inquiry* 16: 417-443.
- Clements, J.C. 1996. *The genesis of a language: the formation and development of Korlai Portuguese*. Amsterdam: John Benjamins.
- Collins, C. 1997. *Local economy*. Cambridge, Mass.: MIT Press.
- Corne, C. 1977. *Seychelles Creole grammar*. Tübingen: TBL Verlag Gunter Narr.
- DeGraff, M. 1997. "Verb syntax in, and beyond, creolization". In Haegeman, 64-94.
- Dowty, D.R., Walls, R.E. and Peters, S. 1981. *Introduction to Montague semantics*. Dordrecht: Reidel.
- Ehrhart, S. 1993. *Le créole français de St Louis (le tayo) en Nouvelle-Calédonie*. Hamburg: Helmut Buske Verlag.
- Ferguson, C.A. and DeBose, C.E. 1977. "Simplified registers, broken language, and pidginization". In *Pidgin and creole linguistics*, A. Valdman (ed.), 99-125. Bloomington: Indiana University Press.
- Foley, W.A. 1986. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
- Foley, W.A. 1997. *Anthropological linguistics*. Cambridge: Cambridge University Press.
- Gadeli, K.E. 1997. *Lesser Antillean French Creole and universal grammar*. Göteborg University PhD dissertation.
- Grimshaw, J. 1997. "Projections, heads, and optimality". *Linguistic Inquiry* 28:372-422.
- Haegeman, L. (ed.) 1997. *The new comparative syntax*. London: Longman
- Hale, K. and Keyser, S.J. (eds) 1993. *The View From Building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass.: MIT Press.
- Hale, K. and Keyser, S.J. 1993. "On argument structure and the lexical expression of syntactic relations". In Hale and Keyser, 53-109. Cambridge, Mass.: MIT Press.
- Halle, M. and Marantz, A. 1993. "Distributed morphology and the pieces of inflection." In

- Hale and Keyser, 111-76.
- Hancock, I. 1996. "The special case of Arctic pidgins". In *Language contact in the Arctic: Northern pidgins and contact languages*, E.H. Jahr and I. Broch (eds), 15-29. New York: Mouton de Gruyter.
- Huttar, G.L. and Huttar, M.L. 1994. *Ndyuka*. London: Routledge.
- Jackendoff, R. 1997. *The architecture of the language faculty*. Cambridge, Mass.: MIT Press.
- Joseph, F.L. 1988. *La détermination nominale en créole haïtien*. Université Paris 7 PhD dissertation.
- Kayne, R.S. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Kihm, A. 1994. *Kriyol syntax: the Portuguese-based creole language of Guinea-Bissau*. Amsterdam: John Benjamins.
- Kihm, A. 1995. "Tayo, the strange bird from New Caledonia: determiners and tense-aspect in Tayo and their implications for creolization theories". *Journal of Pidgin and Creole Languages* 10:225-52.
- Kitahara, H. 1997. *Elementary operations and optimal derivations*. Cambridge, Mass.: MIT Press.
- Laycock, D.C. and Mühlhäusler, P. 1990. "Language engineering: special languages". In *An Encyclopaedia of Language*, N.E. Collinge (ed.), 843-75. London: Routledge.
- Lefebvre, C. 1986. "Relexification in creole genesis revisited: the case of Haitian Creole". In Muysken and Smith, 279-300.
- Lefebvre, C. and Massam, D. 1988. "Haitian Creole Syntax: A case for DET as head". *Journal of Pidgin and Creole Languages* 3: 213-43.
- Li, Y. 1997. "An optimized universal grammar and biological redundancies". *Linguistic Inquiry* 28:170-8.
- Lyovin, A. 1969. "'The Sino-Russian mixed language in Manchuria' by Alina Jablonska". *University of Hawaii Working Papers in Linguistics* 3:135-64.
- Marantz, A. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In *University of Pennsylvania Working Papers in Linguistics* 4.2, A. Dimitriadis et al. (eds.), 201-225.
- Marantz, A. 1999. "Morphology and syntax: paradigms and the ineffable (the incomprehensible, and the unconstructable)". MS.
- McWhorter, J.H. 1998. "Identifying the creole prototype: vindicating a typological class". *Language* 74: 788-818.
- Moysse-Faurie, C. 1995. *Le xârâcùù, langue de Thio-Canala (Nouvelle-Calédonie): éléments de syntaxe*. Paris: Peeters-SELAF.
- Mufwene, S. 1986. "The universalist and substrate hypotheses complement each other". In Muysken and Smith, 129-62.
- Muysken, P. and Smith, N. (eds) 1986. *Substrata versus universals in creole genesis*. Amsterdam: John Benjamins.
- Pinker, S. 1994. *The language instinct*. London: Penguin Books.
- Pollock, J.-Y. 1989. "Verb movement, universal grammar and the structure of IP". *Linguistic Inquiry* 20:325-65.
- Pustejovsky, J. 1995. *The generative lexicon*. Cambridge, Mass.: MIT Press.
- Sadock, J.M. 1991. *Autolexical syntax: a theory of parallel grammatical representations*.

- Chicago: The University of Chicago Press.
- Seuren, P. and Wekker, H. 1986. “Semantic transparency as a factor in creole genesis”. In Muysken and Smith, 57-70.
- Solà, J. 1996. “Morphology and word order in Germanic languages”. In *Minimal ideas: syntactic studies in the minimalist framework*, W. Abraham, S.D. Epstein, H. Thráinsson, and J.-W. Zwart (eds.), 217-51. Amsterdam: John Benjamins.
- Thomason, S.G. and Kaufman, T. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Weinreich, U., Labov, W. and Herzog, M. 1968. “Empirical foundations for a theory of language change”. In *Directions for historical linguistics*, W.P. Lehmann & Y. Malkiel (eds.), 95-195. Austin: University of Texas Press.
- Zwart, J.-W. 1997. “The Germanic SOV languages and the Universal Base Hypothesis”. In Haegeman, 246-67.

The Origin of the Syntax and Semantics of Property Items in the Surinamese Plantation Creole*

Bettina Migge

1. Introduction

In many creoles, elements like Ndyuka *baala* ‘broad’ can function not only as noun modifiers (1), but also as verbs in their own right (2):

- (1) *Saamaka sitaati, na wan baala sitaati.* (Ndyuka)¹
Saramacca street FOC DET broad street
‘Saramacca street is a broad street.’
- (2) *A sitaati baala.*
DET street broad
‘The street is broad.’

This behavior of adjectives in creoles, relatively unusual viewed from the perspective of European languages, has led to a great deal of controversy in the published literature over the proper syntactic and semantic analysis of these items (Alleyne 1980, 1987; Bailey 1966; Bickerton 1973, 1975; Huttar & Koanting 1993; Huttar & Huttar 1994; Sebba 1986; Seuren 1981, 1986; Voorhoeve 1957, 1962; Winford 1993, 1997).

Relatively little research, however, has been conducted on the origin of their syntactic and semantic properties. Two positions have arisen on their origin: Bickerton (1981) argues that the verbal status of these items stems from a language bioprogram, while Alleyne (1980) and Holm (1988) argue that it resulted from substrate influence. Both sides bring little data to bear on the issue, however. Alleyne (1980: 160-2), for example, presents a brief comparison of adjectives in Afro-American language varieties and the Kwa

subfamily (specifically, Twi). He finds systematic parallels between the two groups of languages in this area of grammar but states that “a penetrating examination” of this question is impossible due to “the absence of adequate descriptions of West African languages” (160).

The aim of this paper is to follow up on Alleyne’s (1980) work and determine the impact the substrate languages had on the formation of this area of grammar in the Surinamese Plantation Creole (SPC). To address these two issues, the study draws on both sociohistorical data on the contact setting and on the syntactic and semantic behavior of adjectival items in the modern descendant of the SPC, specifically the Ndyuka varieties, and its primary substrate. Following Thomason & Kaufman (1988) and Thomason (1993), such an undertaking requires the following types of information: first, a determination of all the languages in the contact setting at the time of the creole’s formation; second, an assessment of the relative impact of these languages’ speakers on the formation of the creole; and third, a specification of the exact organization of this area of grammar in the creole and the languages in the contact setting that served as its sources.

The investigation reveals that the emergence of the syntactic and semantic properties of adjectival items in the SPC was primarily the result of substrate influence.

In the following discussion I refer to so-called adjectival items such as *baala* in Ndyuka as *property items* following Thompson (1988). Property items refer to “properties, qualities or characteristics of referents” (ibid. 167). This term is meant to be neutral with respect to the syntactic category into which such items may fall in a given language.

2. Preliminaries

2.1. *The Characterization of Substrate Influence in this Study*

The present study attempts to investigate the formation of the SPC from the perspective of frameworks developed for the study of contact-induced language change. In the literature on language contact, two basic mechanisms of language change are distinguished: borrowing and interference through shift. The two differ from each other with respect to the direction of transfer, the time required for a change to occur, and the order in which types of linguistic items

are transferred. “Borrowing is the incorporation of foreign features into a group’s native language by speakers of that language” (Thomason & Kaufman 1988: 37): the native language is maintained but changed by the addition of the incorporated features. In borrowing or maintenance settings, (extensive) structural borrowing generally emerges as the result of “several hundred years of intimate contact” (ibid. 41) between the different groups. In these situations speakers first borrow vocabulary items, and only as the intensity of the contact increases will they adopt first phonological features and later morphosyntactic ones.

Interference through shift, by contrast, refers to the incorporation of features from the first language of some group into the language of another group which they are in the process of acquiring. In this process the recipient language is changed by the addition of the incorporated features. In such settings “the interference features will enter the TL [target language] as spoken by the shifting [acquiring] speakers quite rapidly [possibly within one generation] though the adoption of these features by the original TL speakers may take more time” (ibid. 41). Interference from the first language “does not begin with vocabulary: it begins instead with sounds and syntax, and sometimes includes morphology as well before words from the shifting group’s original language appear in the TL” (ibid. 39).

The term *substrate influence* (or *stratum interference through shift* [ibid. 38]) has thus been used in the literature on contact-induced change “in relation to those language contact situations involving typically a politically and/or economically dominant group whose superstrate language has been learned imperfectly by the subordinate or substrate group. It attributes to influence from the latter’s language(s) the features of the contact situation’s new language that deviate from those of the TL” (Mufwene 1990: 2). For the purpose of this study, this ‘traditional’ definition is too restrictive since it assumes that in order for substrate influence to occur, the substrate population has to have a target of learning: i.e. it suggests that the (primary) mechanism that instantiates substrate influence is (interference through) shift and that the linguistic outcome of the contact situation is some restructured version of the targeted variety (Thomason 1993). This scenario may in fact apply to some so-called creoles such as Bajan Creole English which appear to have arisen as a result of shift in the traditional sense. However, creoles such as those of Surinam are not cases of targeted shift in the usual sense. Rather, they involve a much higher degree of substrate retention than found in cases of targeted

shift such as Hiberno-English (Winford 1998).

Substrate influence comes about through the use of first or dominant language strategies in the speech of bi- or multilingual speakers, or language learners when attempting to communicate with others who do not share the same varieties. Once such first language strategies “have been replicated by different speakers, repeated by most of them, and established in the contact situation’s new linguistic system (even as a variable feature), they may be characterized genetically as substrate influence. The latter need not be associated synchronically with multilingual speakers and/or SLA” (Mufwene 1990: 2).

For the purpose of this study, the “traditional” definitions of the two mechanisms of change are also too restrictive since they assume that one language is changed under the influence of another but retains its genetic affiliation. Neither of these premises can be taken as given in the case of the formation of a new contact language. In such a case it seems more appropriate to assume that its creators forge a new linguistic system out of linguistic material accessible to them. In this study, lexical and structural features derived from the first or dominant language of the new language’s creators are referred to as (substrate) retention. Features derived from languages other than those of the new language’s creators, such as the superstrate, are referred to as borrowings. Finally, there are also features which cannot be shown to have been derived from either of these sources. They are usually the result of mechanisms such as regularization, levelling, internally motivated change, etc.

2.2. *The Data*

The data for this study would ideally come from the time period in which the formation of the SPC took place. However, since such data is not available for the SPC and its substrate, it is necessary to resort to the closest possible approximation to this data. Thomason (1993) suggests that “modern languages are readily available for study, and three hundred years is not a very long time in language history, so most structures present now were also present then. Any structures that can be reconstructed for the ancestors of some of the languages in question, such as Proto-Kwa, were almost certainly present in those languages when the PC [the pidgin or creole] emerged if they are currently present in relevant daughter languages” (Thomason 1993: 288-289). The aim of this section is twofold: First, it presents some sociohistorical evidence which shows that the modern Ndyuka varieties are usefully repre-

sentative descendants of the original SPC and that the modern varieties of their primary substrate are similarly useful in the study of contact situations which involved their earlier manifestations. Second, it briefly presents the data used in this study.

The Relationship of the Ndyuka Varieties to the SPC

The Ndyuka cluster consists of four varieties spoken by four politically distinct groups of former maroons, known by the names of their varieties: Aluku-Boni, Aukan, Kwinti, and Paramaccan. These varieties are highly mutually intelligible. The Kwinti reside in western Surinam on the Coppename and Saramacca rivers, but all the other Ndyuka groups are presently settled in eastern Surinam and French Guiana, on the banks and islands of the Marowijne river (Paramaccans and Aukans) and its tributaries, the Tapanahoni (Aukan) and Lawa (Aluku-Boni). The Ndyuka groups were founded by slaves who fled the plantations of coastal Surinam between roughly 1710 and 1750. The Aukan are the largest group, numbering roughly 25,000 people; the Aluku-Boni and Paramaccans each number about 5000, and the Kwinti have been estimated to number about 500.

There are three pieces of sociohistorical evidence suggesting that the modern Ndyuka varieties are direct descendants of the original SPC. First, all the Ndyuka communities were founded by slaves who had spent some time, most likely at least a year, on the coastal plantations of Surinam where they must have acquired some variety of the SPC since it served as the lingua franca among the slaves on the Surinamese plantations. Second, the maroons remained in constant and close contact with the slaves on the plantations during their lengthy period of formation (Hoogbergen 1983), which suggests that their varieties probably did not diverge significantly from those spoken on the plantations. Third, throughout their history they have not been influenced significantly by non-maroons, most until now living in their rain forest villages with few close contacts with members of other groups in multiethnic Surinam and French Guiana.

The Relationship of the Modern Substrate Varieties to their Earlier Forms

Sociohistorical evidence suggests that the Africans brought to Surinam during the formative period of the SPC (1680-1720) came principally from two regions. The Dutch Slave Coast of that period was the coast line of present-day Togo and Benin and about 200 kilometers inland. The Loango region was “the

region just north of the Zaire river, i.e., Cabinda, the coastal regions of Congo and Zaire, and southern Gabon” (Arends 1995: 245). Since before the seventeenth century the Dutch Slave Coast region has been primarily inhabited by the so-called Ajatado people, speakers of the Gbe cluster of languages which is part of the Kwa group of the Niger-Congo family (Bendor-Samuel 1989). The Dutch Loango region is primarily populated by speakers of the Kikongo cluster of languages, part of the Kongo group of the Central Bantu branch of Niger-Congo (Grimes 1988). The present study focuses on Gbe for two reasons: first, the slaves drawn from the Dutch Slave Coast constituted the majority of the slaves brought to Surinam during the formative period of the SPC (1680-1720). Second, detailed linguistic data on Kikongo is presently not readily available.²

The speakers of Gbe (literally ‘language’) are distributed over four modern West African states: Ghana, Togo, Benin, and Nigeria. The term Gbe was proposed by the linguist H. Capo in 1977, and has replaced previous names for this continuum referring to single prominent varieties such as Ewe, Aja, Ajatado, Ewe-Fon, Foja, etc. (Capo 1988: 17-32). The roughly 50 members of the Gbe continuum can be roughly subdivided into five main subclusters on the basis of synchronic phonological and morphological evidence: Vhe, Gen, Aja, Fon, and Phla-Phera (Capo 1988: 90-102); see Table 1.

Pazzi (1979) suggests on the basis of historical research that varieties from all five subclusters were involved in the slave trade as either slaves or slave traders.

Table 1. The five subclusters of the Gbe group of languages based on synchronic phonological data (based on Capo 1988: 95-102)

	Kpando Vhlin Vo Agu	Gbin Tɔwu Kpelen Fodome	Ho Aveno Dayin Wance	Awlan Peci Vɛ Waci
	Gen	Anexɔ	Gliji	Agɔi
Gbe	Aja	Dogbo	Stado	Hwe Sikpi
	Fon	Arɔhen Agbome Gun	Kpase Maxi Weme	
	Phla-Phera	Xwela Xwla	Alada Ayizɔ	Tɔli Kotafon Tɔfin Tsaphe

The following evidence suggests that the modern Gbe varieties are direct and representative descendants of the earlier Gbe varieties: first, the sociolinguistic units that existed between 1680 and 1720 still exist today; second, despite political changes, there have been no major social changes in the area that might have led to significant linguistic changes.

The Ndyuka Data

The data for this study come from recordings of natural conversations and formal elicitations. In the recordings of natural conversations, the aim was to record relatively conservative speech samples, and thus they were carried out with people who are either monolingual in the Aukan (AU) or Paramaccan (PM) varieties of Ndyuka, or have only rudimentary control of other languages such as Sranan, Dutch, or French. This group generally consists of women and older men (45+) with very little or no formal education who go to the coastal towns only relatively rarely, interacting primarily with members of their own community. They have generally not lived outside of their rural communities for extended periods of time and are primarily engaged in subsistence farming.

The recordings were carried out either with only the author (PM 7, 17, 18), with a friend of the author (PM 12), or with both (PM 2, 11; AU 2, 3, 4) present. The author either did not participate in the recordings at all (PM 12), or only marginally interacted with the participants in the recordings (PM 2, 11, 17, 18; AU 2, 3, 4). The Paramaccan recordings were carried out during typical everyday activities of the participants; all the Aukan conversations centered around family matters and were recorded during the author's assistant's regular visits to the homes of his family members in the early morning and at dusk.

Some of the data in this study also come from formal elicitations. All the elicitation sessions were carried out in Ndyuka in semi-formal settings: formal in being one-on-one interactions between the author and a native Ndyuka speaker in which the author asked fairly direct questions about the language, informal in that the author was well acquainted with the informants, partially elicited the data informally during other activities, and generally cast them as sessions in which the author was receiving language coaching from the informant. These three informants were in their early 30s and had received some professional training and/or had finished a secondary school education. The Paramaccan informants have both lived in Paramaribo and/or St. Laurent du Maroni in French Guiana for some time, but have remained well rooted in

their respective local networks. The Aukan informant has lived in Paramaribo for quite some time, but his primary social network is made up mainly of members of the Aukan community (living in Paramaribo).

The Aukan informant and one of the Paramaccan informants speak Dutch relatively fluently and some Sranan, while the other Paramaccan informant speaks Sranan relatively fluently and passable Dutch. However, the language for most of their everyday interactions is Ndyuka, and all were judged reliable speakers by various members of the Ndyuka speech community. The decision to chose “non-traditional” informants for the elicitations was due to the fact that the more traditional members of the community tended to be unfamiliar with this kind of direct questioning method, and thus tended to be uncomfortable with, or unwilling to do, this kind of work.

The Gbe Data for the Linguistic Analysis

The Gbe data also comes from a combination of natural recordings and formal elicitations. Natural recordings were carried out with speakers of Aja, Waci, Maxi, and Xwela who are either monolingual or weakly bilingual, generally with very little or no formal education. They are mostly subsistence farmers who spent their lives in their native communities. They were selected by the field workers from among their own social networks and the recordings were carried out in their native communities. The recordings were all carried out with the field worker present and in interactive mode.

Elicitation sessions were carried out with speakers of Anexo (Gen), Fon (Fon), Wance (Vhe), Maxi (Fon), Waci (Vhe), Aja (Aja), and Xwela (Phla-Phera), selected either by the author (interviews conducted in her home in the US) or by H.B.C. Capo (interviews conducted in Garome, Benin). The elicitations were conducted in English and French with the informants living in the USA and in French with the informants living in Benin. All the Gbe informants were university students and are fluent in at least one European language. The USA based informants used both French and a variety of Gbe with friends and family, and all speak more than one variety of Gbe. The Benin-based informants had been living primarily in Benin’s capital, Cotonou, for several years, generally using their variety of Gbe in private interactions and in their native communities, and generally French in official settings.

3. SPC Formation: Sociohistorical Background

3.1. Contact Setting I (1652-1679): Varieties of English

The information and analysis in this section is derived from Rens (1953), Van Stipriaan (1993) and Arends (1995).

Demographically, the first contact setting was characterized by relatively high ratios of Europeans to Africans and high ratios of old slaves to newly imported slaves.

Ethnolinguistically, the early European population consisted mainly of speakers of English, and after 1665 also speakers of Portuguese and Spanish. During the early 1670s the English speakers were replaced by speakers of Dutch, Swedish, German and other European languages. The Africans were most likely native speakers of Gbe, Kikongo and Akan (Twi, Fante, et al.). The early Europeans and slaves coming from other English possessions must have also spoken pidgin and second-language varieties of English.

Socially, this setting was dominated by the small farming unit. Despite a difference in social status between the European and African servants on the one hand, and these two groups of servants and the owner on the other, this setting conditioned relatively little social distance between its members, as well as between the pre-existing population and the newly arriving slaves, since all the members of the farming unit lived and worked in close proximity to each other.

This setting did not provide the necessary social conditions for the emergence of a “radical” creole. The population had a common means of communication, the second language and pidgin varieties of English. Besides, newcomers had sufficient access to these lingua franca varieties and were likely to have also had sufficient motivation to (fully) acquire them since they were in the minority and were able to enter into close social contacts with their speakers. As typically found in language shift situations, they probably also retained some strategies from their native languages in speaking these second-language and pidgin varieties of English.

3.2. Contact Setting II (1680-1694): The Emergence of the SPC

Demographically, the second contact setting was characterized by a rapid growth of the African population, relatively low ratios of Europeans to Afri-

cans, low ratios of old to newly arriving slaves, and a high attrition rate among the slaves.

Ethnolinguistically, the slaves imported during this period consisted of equal numbers of speakers of Gbe and Kikongo. The Europeans were speakers of various European languages. The newly arriving Africans probably did not speak a second-language variety of English and/or an English pidgin upon arrival. Given the practice of “seasoning”, which introduced new slaves to various aspects of plantation life, new slaves are likely to have been exposed at least superficially to pidgin or second-language varieties by Africans who had arrived prior to 1680.

Socially, this was a typical plantation setting, with a much greater number of manual laborers — African slaves — per agricultural unit than the small farming unit, and with a strict hierarchical organization. This would have led to a considerable reduction in the social and work-related contacts and to a relatively large social distance between the different social groups on the plantation, such as Europeans and Africans or elite slaves and field hands. During this period the elite slaves consisted primarily of old slaves and the field hands of the new arrivals.

This setting provided the necessary conditions for the emergence of a contact language such as a radical creole. First, the majority of the plantation population, the newcomers, did not share a common means of communication since they consisted of members coming from two distinct ethnolinguistic groups. Second, the predominance of the new slaves and their relatively strict social separation from old slaves and Europeans would have considerably reduced their access to, and motivation to (fully) acquire, the second-language and pidgin varieties of English used among old slaves and Europeans. Third, the new slaves had a great need for a common means of communication since, given their work and living conditions, they were in close social and professional contacts with one another. To forge their own means of communication, they must have made use of all the linguistic resources available to them, the primary ones being strategies from their native language(s), varieties of Gbe and Kikongo. Their secondary resources must have been the strategies borrowed from the other varieties they were exposed to such as the second-language and pidgin varieties of English spoken by the old slaves. Natively spoken varieties of English would have not been part of the contact setting since its speakers were no longer present.

This suggests that during this period the SPC was made up of a set of

relatively well-defined contact varieties, characterized primarily by strategies retained from the native languages of the (new) slaves and secondarily by strategies borrowed from the second language and pidgin varieties of English spoken by the old slaves.

3.3. Contact Setting III (1695-1720): The Stabilization of the SPC

The final contact setting had the same social characteristics as the second period, but was demographically much more stable and ethnolinguistically more homogenous than the second period.

The population was demographically relatively stable during this period: the ratios of Europeans to Africans continued to decrease at a much slower rate and the ratios of old slaves to new slaves had greatly increased; it was higher than during the first contact setting due to a continually decreasing attrition rate and to a reduction in the number of slaves imported per decade.

Linguistically, about 70% of all the slaves imported during this period were speakers of Gbe. The old slaves would have been divided between two groups of people. The field hands who had arrived during the second period were native speakers of Gbe and Kikongo. They most likely used the varieties of the SPC which they had created during the second period for most of their everyday interactions. Meanwhile, the elite slaves, many of whom had arrived during the first period, were also native speakers of Gbe and Kikongo, but for most of their everyday interactions they most likely used the second-language and pidgin varieties of English current during the first period. However, it seems highly likely that they were also familiar with varieties of the SPC which would have enabled them to communicate with the first group. The Europeans were speakers of various European languages and possibly also speakers of pidgin varieties of English spoken by the second group of slaves with whom they primarily interacted. Few, if any, native speakers of English were present during this period.

By this time, the majority of the population had a common means of communication, the varieties of the SPC which had emerged during the second contact setting. The newcomers to the setting had sufficient access to the varieties of the SPC and must have also had sufficient motivation to (fully) acquire them, since they were in the minority and were in close social contact with its speakers. The newly arriving slaves during the third period must have acquired the varieties of the SPC from its creators, the field hands from the

second period, since they were in close contact with them. It seems unlikely that the new slaves would have acquired the varieties of English spoken by the slaves from the first period, since they were socially relatively sharply separated from them. This suggests that the newcomers most likely shifted to the varieties of the SPC as their primary means of (interethnic) communication. Assuming that the varieties of the SPC differed according to the first language background of its speakers, the newcomers probably shifted to the varieties used by Gbe-speaking slaves; most of them being Gbe speakers themselves, these varieties would have been the easiest for them to acquire. As typically found in all language shift situations, the new slaves probably also retained strategies from their native languages when using these varieties. This suggests that the relatively strong influence from Gbe on the grammatical system of the SPC and its descendants emerged during the third period.

4. The Syntax and Semantics of Property Items in Ndyuka and Gbe: A Comparison

This section examines and compares the discourse functions property items perform, and the syntactic and semantic properties they exhibit in these functions, in Ndyuka and Gbe. The purpose of this comparison is to determine the degree of similarity between these two languages in this area of grammar, in order to determine the role the Gbe substrate played in the formation of the SPC.

My data show that in both Ndyuka and Gbe property items are multifunctional (Bailey 1966; Alleyne 1980; Huttar & Koanting 1993; Winford 1993) or, as Winford (1997) puts it, display flexible categoriality. They may function as verbs and as attributive adjectives. In these functions they display the syntactic and semantic properties typical of prototypical items from these categories. In addition, both languages also have “true” adjectives.

4.1. Property Items Functioning as Verbs

Most property items in Ndyuka and Gbe may function as predicators in their own right. In this function they display various syntactic and semantic properties typical of verbs in these languages.

Syntactic Properties of Property Items Functioning as Verbs

In the literature on property items in creoles, two positions have emerged with respect to the categorial status of property items functioning as predicators. Seuren (1986) argues that they are not verbs, while others such as Bailey (1966), Alleyne (1980, 1987), Sebba (1982), Huttar & Koanting (1993), and Winford (1993, 1997) maintain that they behave like typical verbs in these languages.³ The latter group have argued in favor of a verbal analysis of property items on the basis of the fact that property items may display one or several properties which are assumed to be typical of verbs; a summary of these properties:

- They can be directly preceded by TMA and negation markers.
- They can be predicate clefted.
- They can be followed by degree adverbs.
- Some may take an object argument, as do prototypical change-of-state verbs.

Below I show that property items functioning as predicators in Ndyuka and Gbe should be analyzed as verbs since they display the above properties, which are also typical of verbs in these languages.

The strongest evidence in favor of a verbal analysis of intransitively used property items in Ndyuka and Gbe is their ability to be directly preceded by all kinds of verbal markers such as tense, mood, and aspect (TMA) and negation markers:

- (3) a. *M., a ede uwii fī-i de mu nati namo.*⁴
 M DET head hair POSS-you there OBL wet necessarily
 ‘M., that hair of yours has to necessarily get wet.’ (Ndyuka, PM 17b)
- b. *En mma be e siki.* (Ndyuka, AU 3a)
 her mother PAST PROG sick
 ‘Her mother was getting sick.’
- c. *Emī-dē-e nɔ nyɔ̃ nɛ mī-dè-è.* (Gbe, Maxi 1)
 we-DET-PL HAB be.good for we-DET-PL
 ‘Some of us are generally good for us.’
- d. *Lanmɛ á kɔ̃n.* (Gbe, Xwela 1)
 body FUT fine
 ‘The body will be fine.’

According to my informants, the insertion of a copula morpheme between the

tense and aspect auxiliaries and the property item renders such constructions unacceptable:⁵

- (4) a. **M., a ede uwii fī-i de mu **de** nati namo.*
COP
b. **En mma be e **de** siki.*
COP
c. **Emī-de-e nō **du** nyō nē mī-dè-è.*
COP
d. **Lànme á **de** kòn.*
COP

A second important piece of evidence in favor of a verbal analysis of property items in Ndyuka and Gbe is their ability to be clefted. Via predicate cleft, a verb may be focused or topicalized by copying it to the left edge of the sentence preceding the subject. In Ndyuka the verbal copy is optionally preceded by the focus/topic marker *[n]a* and in Gbe the copy may optionally be followed by a focus marker (*yi* in Aja, *wɛ* in Maxi, *ye* in Waci, and *eyi/ø* in Xwela). In Ndyuka and Gbe all property items that may function as predicates may be predicate clefted (5) just like typical verbs (6):

- (5) a. A *duungu* ye *duungu*? (Ndyuka, PM 18)
FOC drunk you-IMP drunk
'Are you getting DRUNK?'⁶
- b. *Cúɔ* e *cúɔ* na tɔ loo. (Gbe, Xwela 2)
afraid he afraid for water EMPH
'He is AFRAID of water.'
- (6) a. A *lobi* yo *lobi* a sama. (Ndyuka, PM 11a)
FOC love you.FUT love DET person
'You will LOVE the person.'
- b. *Nɔ* ye e *nɔ* aha le *kutɔnu* ɔcɔ. (Gbe, Waci)
drink FOC she drink alcohol LOC name yesterday
'She DRANK alcohol at Cotonou yesterday.'

Property items in Ndyuka and Gbe may also be modified by adverbs of degree just like typical verbs. In both languages, adverbs of degree, including ideophones, generally follow both property items functioning as predicates (7) and typical verbs (8):

- (7) a. *A pikin moi te.* (Ndyuka, PM 17b)
 DET child nice very
 'The child is very nice.'
 b. *É nā sávɔ gānjí.* (Gbe, Maxi 2)
 it FUT big well
 'It will be quite large.'
- (8) a. *Mi e kosi en te.* (Ndyuka, PM 12a)
 I IMP curse him very
 'I was cursing him a lot.'
 b. *Éyɔ dí Bai gánji.* (Gbe, Xwela 1)
 he resemble B. well
 'He resembled Bai a lot.'

Finally, a number of property items in Ndyuka and Gbe may also take a direct object and function like transitive verbs (9):

- (9) *A be go taku a pikin de.* (Ndyuka, PM 17b)
 she PAST go evil DET child there
 'She went to make that child evil.'
- (10) *È gĩgɛ òkǎ lɔ.* (Gbe, Xwela)
 she short string DET
 'She shortened the string.'

In Ndyuka almost all property items from most of the semantic subgroups proposed by Dixon (1977) may take direct objects (with the main exceptions being the verbs *bun* 'good', *lontu* 'round', and *nyun* 'new' from Dixon's semantic subgroups of value, shape, and age respectively). Meanwhile, the elicitations with speakers of Gbe showed that all or several property items from the semantic subgroups of physical property (*cold*, *dead*), dimension (*big*, *broad*), value (*good*, *evil*), color, and weight (*heavy*, *light*) can take a direct object while those from age (*old*, *new*) and human propensity (*lazy*, *stupid*, *evil*) cannot. (In general it seems that the property items used transitively are those referring to properties perceived by native speakers as acquirable through intervention of another entity, specifically humans.)

Thus in both languages property items display syntactic behavior typical of transitive and intransitive verbs in these languages, and moreover, the cluster of properties characteristic of verbs is the same in both languages. This suggests that the syntactic properties of property items functioning as predica-

tors in the SPC were most likely modeled on those of its Gbe substrate.

Semantic Properties of Property Items Functioning as Intransitive Verbs

In the literature on the semantics of property items in creoles, two main views have arisen. Some researchers (Voorhoeve 1957, 1962; Bickerton 1973, 1975; Waite 1983; Sebba 1986; Arends 1989) claim that in their intransitive usage, property items are essentially stative verbs which express current states, but that they may assume a processual interpretation when preceded by the progressive marker (Seuren 1981, 1986; Kahrel 1987). Other researchers (Alleyne 1980, 1987; Winford 1993, 1997) argue that most property items are non-stative and that their current state reading is ambiguous with a completed process reading.⁷

The latter group of researchers have argued in favor of a non-stative analysis of property items on the basis of the fact that property items display one or several of the following properties:

- They take on an inchoative interpretation when marked by the imperfective aspect marker.
- They have a completed process reading when marked by the completive aspect marker.
- Their current state reading is an entailment of the past action reading.

Below I show that property items functioning as intransitive verbs in Ndyuka and Gbe should be analyzed as non-stative verbs since they display all the properties above, which are typical of non-stative verbs in these languages.

The strongest evidence in favor of the argument that property items in Ndyuka and Gbe are processual or non-stative in character is their ability to take an inchoative interpretation when preceded by the imperfective or progressive aspect marker, i.e. the progressive marker indicates that the subject is in the process of acquiring the property denoted by the property item.

In Gbe all verbal property items modified by the progressive aspect marker take on an inchoative interpretation. In Ndyuka there is a small subset of items which generally do not take on an inchoative meaning when they are preceded by the progressive aspect marker *e*. In Ndyuka *e* directly precedes the verb (11a). In some of the Gbe varieties such as Aja, Maxi, and Waci, intransitive verbs marked for progressive aspect are reduplicated as well as preceded by the non-nominal copula (*le* in Aja, *ɖu* in Maxi, and *lɛ* in Waci) and followed by a particle ([*k*]ɔ in Aja, *we* in Maxi and Fon, *ŋ* in Waci) (11b).⁸

In other varieties of Gbe such as Gen and Xwela, the progressive marker, *le* and *nɔ* respectively, simply precedes the unreduplicated verb (11c):

- (11) a. *Sama abi a wasi duku di e* (Ndyuka, PM 12a)
 person have DET wash towel REL PROG
nati ape?
 wet there
 ‘Who has the towel that is getting wet there?’
- b. *Nudé ká **ɖo** nyu-nyɔ ne mɛ weǎ.* (Gbe, Maxi 2)
 nothing ? COP good-good for person FOC
 ‘Nothing has even been getting good/better for people.’
- c. *Emɛ lɔ **nɔ** ga.* (Gbe, Xwela)
 person DET PROG big/fat
 ‘The person is getting fat.’

Further evidence of the non-stative or processual semantics of property items functioning as intransitive verbs can be derived from the fact that most property items may have a past action or completed process reading when they are modified by the VP-final completive aspect marker, *kaba* in Ndyuka (12a) and in Gbe, *vɔ* in all dialects, with the additional alternate *fo* in Maxi and Fon (12b):

- (12) a. *Nounou a **dee** kaba.* (Ndyuka, PM 17a)
 now it dry COMP
 ‘Now it [her breast] has gotten dry/is already dry.’
- b. *Awu lɔ **fa** vɔ.* (Gbe, Aja)
 shirt DET wet COMP
 ‘The shirt has gotten wet/is already wet.’

The past action or completed process interpretation of property items is not restricted to property items marked by the completive aspect marker, however; property items not modified by the completive aspect marker may also have a completed process reading besides a current state interpretation (13). In the case of constructions without overt TMA marking, context determines whether they have a stative (14a-b) or a non-stative (14c-d) reading:

- (13) a. *A man dede.* (Ndyuka)
 DET man dead
 ‘The man has died/is dead.’

- b. Èé có kú. (Gbe, Maxi 2)
 they all dead
 ‘They have all died/are all dead.’
- (14) a. Di mi doo osu a man be dede. (Ndyuka)
 REL I arrive house DET man PAST dead
 ‘When I came home, the man was dead.’
- b. Mε dehwlē le kū gbón é mɔ bayi nu tetebi
 person certain PL die till they PAST make thing all
 nε.
 for.them
 ‘Certain people were dead and they had performed all the rituals for them.’ (Gbe, Xwela 1)
- c. Te den kaba feti wan pisi [...] neen (Ndyuka, AU 2a)
 when they finish fight one piece then
 a dede.
 he die
 ‘When they had finished fighting a bit, then he died.’
- d. Dawūé ɔ mɔ jε aɣwā, mɔ kū xwi. (Gbe, Xwela 1)
 man DET PAST fall PAST die IDEO
 ‘The man fell and died quietly.’

The above examples suggest that most verbal property items in Ndyuka and Gbe are basically non-stative, process-denoting verbs which can take on a stative reading given the right context. The stative reading of such items is always “an entailment of the completed process reading, and does not represent a distinct syntactic function” (Winford 1997: 263).

In Ndyuka and Gbe there are a small number of property items which are more stative in nature.⁹ As unmarked verbs they have a stative interpretation “without implying a past process” (Winford 1997: 26) (15a-b). In both languages such items come from the semantic subgroups of human propensity and value. In Ndyuka they may not be marked for progressive aspect (15c) and in both Ndyuka and Gbe they do not allow a past action or completed process reading when marked by the completive aspect marker. Instead, they seem to convey that the state they denote has already been in existence for some time (15d):

- (15) a. *A telefon a bunkopu.* (Ndyuka, AU 1)
 DET telephone NEG cheap
 ‘The telephone is not cheap/*has not gotten cheap.’
- b. *Awu lo vè xo.* (Gbe, Aja)
 clothe DET expensive market
 ‘The piece of clothing is expensive/*has gotten expensive.’
- c. **A telefon e bunkopu.* (Ndyuka)
 DET telephone PROG cheap
- d. *A telefon bunkopu kaba.* (Ndyuka)
 DET telephone cheap COMP
 ‘The telephone is already cheap/*has already gotten cheap.’

These semantic similarities between Ndyuka and Gbe suggest that the semantic properties of property items functioning as intransitive predicators in the SPC were most likely modeled on those of Gbe.

4.2. Property Items Functioning as Predicative Adjectives

This section investigates property items which occur in predicate position but which cannot function as verbal heads in their own right. In both Ndyuka and Gbe these property items are typically complements to a copula. There are two types of predicative adjectives, unreduplicated and reduplicated property items.

Unreduplicated Property Items Functioning as Predicative Adjectives

In Ndyuka there are not any unreduplicated property items which categorically require the copula *de* when they occur in predicative position.¹⁰ There are, however, a small number of items such as *bun* ‘good, well’, *moi[n]* ‘nice, well’, *nyun* ‘new’, *fanya* ‘disorganized’, and *pii* ‘quiet’ which may be complements to *de* when occurring in predicative position: compare (16a) and (16b). When functioning as predicative adjectives as in (16a), such property items are clearly not verbal:

- (16) a. *Ma i libi an de bun.* (Ndyuka, PM 7b)
 but your life NEG COP good
 ‘But your life is not in a good state.’
- b. *A kii feti fu den, a an bun.* (Ndyuka, PM 17a)
 DET kill fight POSS them it NEG good
 ‘Their violent fights, it is not good.’

On close investigation it becomes clear, however, that the semantics of *bun* functioning as a predicative adjective (16a) and that of verbal *bun* (16b) are not completely equivalent. In (16b) *bun* describes a more permanent property of its referent, *a kii feti* ‘the violent fights’. In (16a), on the contrary, *bun* describes a temporary state in which its referent *i libi* ‘your life’ exists at this point in time.¹¹ The semantics of (16a) is equivalent to similar constructions involving the reduplicated manifestation:

- (17) *Efu den sikin de bunbun da a bun!* (Ndyuka, AU 2a)
 if their body COP good then it good
 ‘If their bodies are in a good/healthy state, then it is OK.’

The semantic similarities between the reduplicated and the unreduplicated property item suggests that the unreduplicated property item in (16a) is a shortened form of the reduplicated property item in (17), as proposed by two of my informants. Such an analysis might be supported by the observation that constructions of the sort exemplified in (16a) are only available for a few items in Ndyuka, most of which are in frequent usage. Temporary states are typically expressed using the construction exemplified in (17) in the Surinamese creoles (see below for further discussion).

In Gbe a few property items from several semantic subgroups such as age, shape, color and dimension are always complements to the non-nominal copula (*le* in Aja, *ɔ/u* in Maxi, *lɔ* in Waci and *ɔ* in Xwela) in predicative position:

- (18) a. *Mɔ̀tò ɔ ɖù yōyɔ́.* (Gbe, Maxi)
 car DET COP new
 ‘The car is new.’
 b. *Fí ā lɔ̀ klòbòtò.* (Gbe, Waci)
 place DET COP round
 ‘The place is round.’
 c. *Elan lɔ̀ le mumu.* (Gbe, Aja)
 meat DET COP raw
 ‘The meat is raw.’

As in Ndyuka, property items functioning as complements to the copula in Gbe are not verbal.

Reduplicated Property Items Functioning as Predicative Adjectives

In both the Surinamese creoles and Gbe, verbal property items from all semantic subgroups and change-of-state verbs may be reduplicated to express temporary and visible states. The reduplicated property items are not verbal since they are always complements to the non-equative copula (19a, b) or other state-denoting verbs (19c, d) when occurring in predicative position. They are best classified as predicative adjectives:¹³

- (19) a. *Duunguduungu a i de.*¹⁴ (Ndyuka, AU4b)
 drunk-drunk he HAB COP
 'He is usually (IN A) DRUNKEN (STATE).'
- b. *Awu ɔ ɖu fi-fa (ji).* (Gbe, Maxi)
 clothe DET COP wet-wet on
 'The shirt is (in a) wet (state).'
- c. *I bee tan deede fu te i dede.* (Ndyuka, PM 17b)
 your belly stay dry for until you die
 'You will remain (in an) infertile (state) until you die.'
- d. *I mɔ mɔto ɔ (ɖe) gbi-gba.* (Gbe, Xwela)
 I see car DET COP break-break
 'I found the car (in a) broken (state).'

Constructions involving reduplicated adjectives expressing temporary and visible states are found in contexts referring to:

The continuity of a state contrary to expectation (the man was quite old):¹⁵

- (20) a. *Ma ala en ede be de baakabaaka.* (Ndyuka, PM 11b)
 but all his head PAST COP black-black
 'But all his head (hair) was still (in a) black (state).'

The continuity of a state over time:

- b. *Yé du wɔ ɸodiɸodí mɔ a nɔ ɸodiɸodí ji.*¹⁶
 COR village PL dirty-dirty in FUT stay/COP dirty -dirty on
 'And your dirty village will always be in a dirty state.' (Gbe, Waci 2)

Ignorance of the process or the agent that caused the state:

- c. *Amamεε kpo ó vá yí kpɔ é.* (Gbe, Waci 1)
 naked-naked only they come go see him
 'They found him [in a] totally naked [state].'

A temporary state:

- d. *A di den tu man de be de nyoninyoni.*
 FOC REL DET-PL two man there PAST COP small
 ‘It was when the two there were still [in a] young [state].’
 (Ndyuka, PM 17b)

As observed by Huttar & Huttar (1997), reduplicated property items in Ndyuka may also express approximation (21a) or distribution (21b). Like their unreduplicated counterparts, they are, however, fully verbal since they can be directly modified by TMA markers (as in 21a), be predicate clefted (21c), and may also function as transitive verbs (21d):

- (21) a. *Sensten di a pikin de, de e nyoninyoni.*
 since REL DET child there they PROG small-small
 ‘Since the child is there, they [the breasts] are always smallish.’
 (Ndyuka, PM 17b)
- b. *En ede piipii moo du fi i.* (Ndyuka)
 his head peel-peel more one POSS you
 ‘His head has gotten more bald spots than yours.’
- c. *Na natinati den natinati a osu.* (Ndyuka)
 FOC wet-wet they wet-wet DET house
 ‘They made the house KIND OF WET.’
- d. *Den baakabaaka mi buuku.* (Ndyuka)
 they black-black my trousers
 ‘They have made my trousers blackish.’

Semantically, they are also similar to their unreduplicated counterparts: they are processual, and their current state reading is an entailment of their (in)-complete past process reading.

The set of reduplicated property items in Ndyuka which may express approximation is more restricted than that which expresses temporary and visible states. It is rather difficult, however, to determine the exact group of property items which may be reduplicated to express approximation or distribution. Generally, it seems that only those property items denoting properties which the native speaker can imagine assuming an intermediate quality can undergo this process of reduplication. (This type of reduplication seems to be also widespread in Sranan [Winford 1997], but is uncommon in Saramaccan according to my Saramaccan informant.)¹⁷

The above discussion reveals the following similarities between Ndyuka (and the Surinamese creoles) and Gbe: both have reduplicated and unreduplicated property items which function as predicative adjectives in predicative position; in both language groups they are complements to the non-equative copula (or other state-denoting verbs in the case of reduplicated property items); in both, the number of unreduplicated property items functioning as predicative adjectives are very few in number; in both, reduplicated property items express temporary and visible states. The above similarities suggest that the syntax and semantics of predicative adjectives in the SPC were most likely modeled on that of the Gbe varieties.

4.3. Property Items Functioning as Attributive Adjectives

Previous studies of property items in creoles (Bailey 1966; Alleyne 1980; Huttar & Koanting 1993; Winford 1993, 1997; Huttar & Huttar 1994) have shown that they may also function as attributive adjectives. In this function they usually have the same form as in predicative position and are found to exhibit the following cluster of properties:

- They may modify nouns.
- They may be conjoined with one or more adjectives in the NP.
- They may be modified by adverbs of degree in the NP.

Below I show that according to the above properties, property items in attributive position in both Ndyuka and Gbe behave like typical predicative adjectives.

The strongest evidence that property items in Ndyuka and Gbe behave like typical attributive adjectives is their ability to function as modifiers of the nominal head of an NP. In the Surinamese creoles, including Ndyuka, property items functioning as noun modifiers precede the nominal head (23a, b) while in Gbe they follow it (23c, d):¹⁸

- (22) a. *Mi si a-i dongo e pasa anga a siki pikin.*
 I see he-PROG go.down PROG pass with DET sick child
 'I saw him going down river with the sick child.' (Ndyuka, AU 3a)
- b. *A booko di lontulontu tafa.* (Saramaccan)
 he break DET round-round table
 'He broke the round table.'

- (23) a. *Xi éxé mé kūku lɛ nɔ dɛ kūtónú ɔ.*
 place REL person dead-dead PL PR LOC name DET
 ‘...the place where the corpses are at Cotonou.’ (Gbe, Xwela 1)
- b. *Emē ɖaxó elimɔ́ wá tɔ ɔ nũ.* (Gbe, Maxi 1)
 person old Elimon come water DET mouth
 ‘The elder Elimon came to the well.’

Note also that in Saramaccan and the Gbe varieties most property items functioning as attributive adjectives are reduplicated while in Ndyuka (and Sranan) they are not.

Like typical adjectives, property items in both language groups may also be conjoined in the noun phrase. They are typically juxtaposed without a conjunction:

- (24) a. *Wan gaan bigi fatu mabee.* (Ndyuka, PM 17a)
 DET very big fat type of fruit
 ‘a very big and fat mabee.’
- b. *Yi awu gbo wewei lo na.* (Gbe, Xwela)
 take shirt big white-white DET give.me
 ‘give me the big white shirt.’

Finally, property items functioning as noun modifiers can be modified by adverbs, suggesting that they are the head of an adjectival phrase. In Ndyuka adverbial modifiers generally precede property items in an adjectival phrase (24a, 25a) but ideophones functioning as adverbial modifiers generally follow it (26a). In Gbe both kinds of modifiers follow the adjective:

- (25) a. *Wan gaan langa pisi beele N. langa gi en.*
 DET very long piece bread N. hand give her
 ‘N. handed her a very long piece of bread.’ (Ndyuka, PM 12a)
- b. *Emē zɪza gāji.* (Gbe, Xwela)
 person long-long very
 ‘A very tall person’
- (26) a. *Wan deki guguguu wan.* (Ndyuka, OK 4b)
 DET fat IDEO DET
 ‘A very fat one (person).’
- b. *Nya gli kpowunkpowun ce nɔhun.* (Gbe, Maxi 3)
 man short IDEO my similar
 ‘A very short man like me.’

The above discussion shows that property items in the Surinamese creoles and Gbe display the same behavior when they occur in attributive position and are clearly adjectives. The behavior of attributive adjectives in Ndyuka (and the Surinamese creoles), however, differs from that of Gbe with respect to their constituency ordering in the NP.

4.4. *The Relationship between Property Items in Attributive and Predicative Position*

In Ndyuka most property items that may function as verbs in their own right, including reduplicated property items expressing approximation or distribution, may also occur in attributive position without undergoing a change in form. In (27) *tyali* and *gaandigaandi* function as attributive adjectives, and in (28) as verbs:

- (27) a. *Ai baala, **tyali** kondee, **tyali** goontapu.* (Ndyuka, AU 3b)
 yes brother sad country sad world
 ‘Yes brother, a sad country/village, a sad world.’
- b. *Ne en **gaandigaandi** yuu a kon ya.*
 LOC his old-old hour he come here
 ‘During his middle aged years he came here.’ (Ndyuka, PM 11b)
- (28) a. *Da i sa **tyali** wan dei gi a sama namo.*
 then you can sad one day give DET person necessarily
 ‘Then you could definitely mourn one day for the person.’
 (Ndyuka, AU 3b)
- b. *A man ya **gaandigaandi**.* (Ndyuka)
 DET man here old-old
 ‘The man is middle-aged.’

Some property items, which have traditionally been categorized as change-of-state verbs, such as *faa* ‘cut tree’, *doo* ‘split’, *tei* ‘tie’, *boli* ‘cook’ and *fon* ‘pound’ must always be reduplicated in attributive position with the unreduplicated form apparently unacceptable:

- (29) *Gi mi a *fon/fonfon alisi.* (Ndyuka)
 give me DET pounded rice
 ‘Give me the pounded rice.’

Property items such as *nyun* ‘new’ and *libi* ‘alive’ may appear in both simple and reduplicated form in attributive position without an apparent distinction in meaning:¹⁹

- (30) *Den go bai wan nyun/nyunyun masini.* (Ndyuka)
 they go buy DET new machine
 ‘They went to buy a new out-board motor.’

Reduplicated property items which express temporary and visible states, and in predicative position are not verbal but are always complements to the non-equative copula or other state-denoting verbs (see examples 19 and 20 above), generally cannot appear in attributive position. In attributive position, these reduplicated property items generally have an approximative interpretation, as suggested by the B reading in (31b):

- (31) a. *Ma ala en ede be de baakabaaka.* (Ndyuka, PM 11b)
 but all his head PAST COP black-black
 ‘But all his head (hair) was still (in a) black (state).’
 b. *A gi mi wan baakabaaka impi.* (Ndyuka)
 he give me DET black-black shirt
 *A: ‘He gave me a black shirt.’
 B: ‘He gave me a blackish shirt.’

Finally, two property items in Ndyuka, notably *gaan* ‘great’ and *pikin* ‘small’, may only occur in attributive position (32a). They cannot occur in predicative position as either predicators in their own right or as complements to the copula *de* (32b): in predicative function *gaan* is replaced by *bigi* ‘great, big’ and *pikin* is replaced by *nyoni* ‘small’ (32c):

- (32) a. *Da pe a gaan nefi de?* (Ndyuka, PM 2a)
 then where DET big knife COP
 ‘And where is the big knife?’
 b. **A nefi (de) gaan.* (Ndyuka)
 DET knife COP big.
 c. *A nefi bigi.*
 DET knife big.
 ‘The knife is big.’

It is not clear to me at this point, however, whether the relationships that exist between property items in predicative and attributive position in Ndyuka

discussed above are entirely representative of those that existed in the SPC, since according to my informant work on Saramaccan they are not entirely the same in that language.

In Gbe most property items which may function as verbs are always reduplicated when occurring in attributive position:

- (33) a. *E dɔ asán tantɔn di ɔn é na xú.* (Gbe, Maxi 3)
 they say day eight now DET it FUT dry
 ‘They say in eight days now it will be dry.’
 b. *Emí ba agbɔ xúxú de ǎ.* (Gbe, Maxi 3)
 we want vegetable dry-dry DET NEG
 ‘We don’t want this kind of dry vegetables.’

A much smaller number of verbal property items and predicative adjectives do not undergo a change in form when functioning as attributive adjectives:

- (34) a. *Imɔ lɔ kike.* (Gbe, Xwela)
 street DET wide
 ‘The street is wide.’
 b. *Yi nɔ imɔ kike ji* (Gbe, Xwela 1)
 I live street wide on
 ‘I live on a wide street.’

Finally, some items may never occur in predicative position:

- (35) a. *Mía gbá nyí agbá gan de.* (Gbe, Waci 1)
 your load COP load great one
 ‘Your load is a big one.’
 b. **Mia gba [lɔ] gan.* (Gbe, Waci)
 your load COP great

Some reduplicated and very few unreduplicated property items in Gbe may take the suffix *-(y/d)i* when occurring in attributive position (36). It is not clear to me at this point whether its use is obligatory with (some) property items and whether it is a derivational morpheme:

- (36) a. *Yée bó hwɛhwɛyí.* (Gbe, Xwela 2)
 their fetish small
 ‘their small fetish.’
 b. *Emɛ gẽgẽí.* (Gbe, Xwela)
 person short
 ‘a short person’

The above discussion reveals the following similarities between Ndyuka and Gbe. First, in both language groups there are (i) property items that may function as both attributive adjectives and verbs without undergoing a change in form, (ii) property items that are unreduplicated when functioning as verbs but are reduplicated when functioning as attributive adjectives, and (iii) property items that may only occur in attributive position. There are, however, also a number of differences between Ndyuka and Gbe in this area: in Ndyuka there are also property items which are found in both reduplicated and unreduplicated form in attributive position; in Gbe, most verbal property items of all semantic types are reduplicated in attributive position while in Ndyuka they usually do not undergo a change in form; in Gbe reduplicated property items functioning as predicative adjectives (expressing visible and temporary states) may also occur in attributive position, but not in Ndyuka. These differences between Ndyuka and Gbe suggest that the Gbe languages did not serve as the (only) model for this area of grammar in the SPC. The differences between Ndyuka and Saramaccan in this area of grammar suggest that there was variation in this area of grammar in the SPC, most likely originating from the differences that existed in this area of grammar between the languages that served as inputs to the formation of the SPC.

5. Summary: Property Items in Ndyuka and Gbe

The comparison of the behavior of property items in Ndyuka and in its Gbe substrate reveal properties shared by both languages and those not shared by them. Below I briefly summarize these properties.

Properties shared:

1. Property items functioning as predicates exhibit the same behavior as typical verbs in these languages:
 - (a) they may be marked for tense, mood, and aspect
 - (b) they can be predicate clefted
 - (c) they may be marked by adverbial modifiers.
2. Some property items may also function as transitive verbs.
3. Most property items functioning as intransitive verbs seem to have a non-stative and or processual interpretation:

- (a) they take on an inchoative interpretation when marked by the progressive marker
 - (b) they take on a completed process reading when marked by the completive marker.
 - (c) their current state reading is an entailment of their past action or completed process reading.
4. A small number of verbal property items seem to be inherently stative or non- processual since their current state reading does not seem to imply the existence of a past process.
 5. There are reduplicated property items which are always complements to the non-equative copula in predicative position.
 6. Reduplicated property items which are always complements to the non-equative copula in predicative position express temporary and visible states.
 7. Property items functioning as attributive adjectives exhibit the following behavior typical of adjectives:
 - (a) they modify nouns
 - (b) they may be conjoined in the NP
 - (c) they may be modified by adverbs of degree.
 8. There are several relationships between property items in predicative and attributive position:
 - (a) some property items have the same shape in attributive and predicative function
 - (b) some property items are reduplicated in attributive function
 - (c) some property items can only occur in attributive function.

Properties not shared:

1. In attributive position:
 - (a) Most property items in Gbe are reduplicated while in Ndyuka they are not.
 - (b) Some property items in Ndyuka may be reduplicated or unreduplicated in attributive position.
 - (c) Most reduplicated property items expressing temporary and visible states may not function as attributive adjectives in Ndyuka.

- (d) In Gbe property items functioning as noun modifiers typically follow the noun they modify while in Ndyuka they precede it.
- 2. In Ndyuka property items can be reduplicated to express approximation or distribution. They function like typical verbs in predicative position and like adjectives in attributive position.

The results of the comparative study suggest that the behavior of property items in the SPC was modeled primarily on that of property items in Gbe, since they share a great number of important similarities in both languages. The evidence from the Surinamese creoles taken together (Ndyuka, Saramaccan, and Sranan) suggests, however, that there might have been variation in the (early) SPC in the area of nominal attribution. Finally, the study also suggests three possible areas where the pidgin and second-language varieties of English from the first contact setting might have influenced the grammar of the SPC: first, these varieties may have been (one of) the sources for unreduplicated property items functioning as attributive adjectives; second, they may have determined the constituent order within the NP between nouns and their various modifiers; third, they were of course the source for the lexical items, specifically their phonological shape and possibly some of their syntactic and semantic features.

6. Conclusion: Implications for Theories of Creole Formation

These findings confirm theories maintaining that creole formation is the result of contact-induced language creation in which several processes such as substrate retention, borrowing, simplification, etc. were operative (Mufwene 1990, McWhorter 1997, Winford 1998). The evidence suggests that the formation of the SPC involved two mechanisms: the retention from the primary substrate of the syntactic and semantic behavior of property items on one hand, and on the other hand the adoption of the phonological shapes of property items, the constituent order within the NP, and possibly one of the strategies for deriving attributive adjectives from verbal property items from second-language and pidgin varieties of English.

This study challenges the view espoused by those characterizing substrate influence in creole formation as taking place within targeted language shift (Thomason & Kaufman 1988). The sociohistorical background to the formation of the SPC made it very unlikely that the creators were shifting to

the superstrate varieties, given how very little contact they must have had with their speakers, and the low possibility of acquiring that would have resulted.

The evidence also challenges the bioprogram hypothesis (Bickerton 1981, 1984, 1989; Veenstra 1996) which maintains that the highly multilingual contact settings of creole genesis forced slaves' children to develop a viable language via recourse to the blueprint of language presumably innate to the human species. The sociohistorical evidence shows that during the formation period of the SPC, most slaves would have had sufficient access to speakers of their native language and enough knowledge of varieties (of English) for some communication with their superiors, the elite slaves and administrators. In their development of a new language for interethnic communication, the field hands thus would have made recourse to these varieties rather than an innate language blue-print.

Notes

- * An earlier version of this paper was presented at the Annual Meeting of the Society for Pidgin and Creole Linguistics (Chicago, January 1997). I would like to thank D. Winford for comments and criticisms on several aspects of this paper. Special thanks are due to my informants J.-P. Agbogba, K. D. Agbemebia, J. Asodanoe, F.-X. Atchade, U. Deel, D. Gagnon, M. Gbenado, N. Y. Houngbo, S. B. Sana, C. Velanti, and F. Velanti, who kindly supplied the elicited data for this study. I also express my sincere appreciation to all the people who willingly participated in the recordings. The informant work was generously funded by the Department of Linguistics of the Ohio State University. Any errors and omissions are, of course, the responsibility of the author.

1. Unless otherwise indicated, the data was collected by the author.
2. A reviewer suggested that the fact that there are lexical items derived from Kikongo in all of the Suriname creoles is evidence that Kikongo speakers had a major influence on the formation of the SPC. I do not wish to deny any influence from Kikongo on the formation of the SPC; however, I seriously doubt that the existence of lexical items by themselves demonstrates it, and the importation figures suggest that the Gbe varieties most likely played the more important role. In the meantime, I have been unable to locate speakers of Kikongo up to this writing, and the few grammars contain little data or analysis of property items and have been rated in general as unreliable (Salikoko Mufwene, personal communication, March 1998).
3. For a complete discussion of the controversy, see Winford (1997).
4. Examples drawn from the natural conversations contain a tape number. Examples without a tape number come from elicitations.
5. Ndyuka has only one true copula morpheme, namely *de*, which can take as complements a wide variety of non-verbal items including some adjectives (see below). Other items

which have been assumed to be copulas, such as *kon* and *tan*, in fact are not. They are apparently both main verbs with aspectual meaning, but do not occupy AUX or INFL. When *kon* 'come' precedes a verb it denotes that the state or action denoted by it has been accomplished. *Tan* 'to remain, to stay' may function as a marker of continuity or stativity when preceding property items; i.e. it expresses that the property denoted by the property item has been in existence for quite some time. The element *toon* typically subcategorizes for NPs which denote a title or job only acquirable through learning, study, etc.

Most varieties of Gbe have two copulas, namely *de*, *dɔ/u*, *lɔ*, *le* and *nyi*. *Nyi* is only used for nominal predication while *de*, *dɔ/u*, *lɔ*, and *le* may be used to predicate a wide variety of non-verbal items including adjectives.

6. Capitalization in the translation indicates focus or emphasis.
7. For a complete discussion of the controversy, see Winford (1997: 259-280).
8. Transitive verbs marked for progressive aspect are not reduplicated, but the direct object and the verb switch places.
9. See also Huttar & Koanting (1993) and Huttar & Huttar (1994) for a similar analysis with respect to Ndyuka.
10. Exceptions to this are ideophones which denote properties; because their categorial analysis is currently controversial, they will not be included in this discussion.
11. See Arends (1989) for a similar observation regarding Sranan.
12. In the Surinamese creoles and in some Gbe varieties (Waci, Gen, Aja), the whole word is reduplicated. In other varieties (Maxi, Fon, Xwela) only the initial consonant is reduplicated and followed by the epenthetic vowel. Note also that reduplicated and unreduplicated property items in Gbe may be followed by the adjectival suffix *i* (Capo 1983). It seems to be obligatory with some items and optional with others.
13. See Huttar & Koanting (1993) and Huttar & Huttar (1994) for a similar observation.
14. In this example, the adjective has been topicalized.
15. In (urban varieties of) Sranan, reduplicated property items which are complements to *de* seem to only express temporary and visible states in context (20a).
16. In non-present tense constructions, the non-nominal copula *lɔ* is replaced by *no* in Waci.
17. My own informant work as well as Winford (1997) suggest that in (urban) Sranan, reduplicated property items which are complements to *de* may also (and more readily) express approximation.
18. It is not quite clear why attributive adjectives occur prenominally in the Surinamese creoles and postnominally in their substrate, including Kikongo. A possible explanation could be that the Surinamese languages inherited the constituent order within the NP from the predecessor pidgin or second-language varieties of English which served as inputs to the formation of the SPC.
19. Some property items such as *siki* and *lau* have different meanings in the simple and in the reduplicated form. *Lau*, for example, means 'crazy' and *laulau* means 'bad' in attributive position (Huttar & Huttar 1997).

References

- Alleyne, M. C. 1980. *Comparative Afro-American*. Ann Arbor: Karoma.
- Alleyne, M. C. 1987. "Predicate structures in Saramaccan". In Alleyne, 71-88.
- Alleyne, M.C. (ed.) 1987. *Studies in Saramaccan language structure*. Mona, Jamaica: Folklore Studies Project, University of the West Indies.
- Arends, J. 1989. *Syntactic developments in Sranan*. Catholic University of Nijmegen Ph.D. dissertation.
- Arends, J. 1995. "Demographic factors in the formation of Sranan". In *The early stages of creolization*, J. Arends (ed.), 233-285. Amsterdam: John Benjamins.
- Bailey, B. 1966. *Jamaican creole syntax*. Cambridge: Cambridge University Press.
- Bendor-Samuel, J. 1989. *The Niger-Congo languages*. New York: University Press of America.
- Bickerton, D. 1973. "On the nature of a creole continuum". *Language* 49: 640-669.
- Bickerton, D. 1975. *Dynamics of a creole system*. Cambridge: Cambridge University Press.
- Bickerton, D. 1981. *Roots of language*. Ann Arbor: Karoma.
- Capo, H. B. C. 1983. "'I' et 'U' en Hwe et leur place dans la reconstruction du Proto-Gbe". *Journal of West African Languages* 13: 3-18.
- Capo, H. B. C. 1988. *Renaissance du Gbe: Reflexions critiques et constructives sur l'Eve, le Fon, le Gen, l'Adja, le Gun, etc.* Hamburg: Helmut Buske.
- Dixon, R. M. W. 1977. "Where have all the adjectives gone?" *Studies in Language* 11: 19-80.
- Grimes, B. 1988. *Ethnologue: Languages of the world*. Dallas: SIL.
- Holm, J. 1988. *Pidgins and creoles*, Vol. I. Cambridge: Cambridge University Press.
- Hoogbergen, W. 1983. "Marronage en marrons, 1760-1863: De niet-gepacificeerde marrons van Suriname". In *Suriname, de schele onafhankelijkheid*, G. Willemsen (ed.), 75-110. Amsterdam: de Arbeiderspers.
- Huttar, G. L. and Huttar, M. L. 1994. *Ndyuka*. London: Routledge.
- Huttar, G. and Huttar, M. 1997. "Reduplication in Ndyuka". In *Pidgins and creoles: Structure and status*, A. Spears and D. Winford (eds), 395-414. Amsterdam: John Benjamins.
- Huttar, G. and Koanting, E. 1993. "Verbs and adjectives: Categories or a continuum?" Paper presented at the annual meeting of the Society for Pidgin and Creole Linguistics, Los Angeles.
- Kahrel, P. 1987. "Stative verb formation in Saramaccan". In Alleyne, 71-88.
- McWhorter, J. 1997. *Towards a new model of creole genesis*. New York: Peter Lang.
- Mufwene, S. 1990. "Transfer and the substrate hypothesis in creolistics". *Studies in Second Language Acquisition* 12: 1-23.
- Pazzi, R. 1979. *Introduction à l'histoire de l'aire culturelle ajatado*. (Études et Documents en Sciences Humaines no. 1.) Institut National des Sciences de l'Éducation, Université du Bénin, Lomé, Togo.
- Rens, L. 1953. *The historical and social background of Surinam's Negro English*. Amsterdam: North-Holland.
- Sebba, M. 1982. "Adjectives and comparatives in Sranan Tongo." *Amsterdam Creole Studies* 4: 1-22.

- Sebba, M. 1986. "Adjectives and copulas in Sranan Tongo". *Journal of Pidgin and Creole Languages* 11: 109-121.
- Seuren, P. 1981. "Tense and aspect in Sranan". *Linguistics* 19: 1043-1076.
- Seuren, P. 1986. "Adjectives as adjectives in Sranan: A reply to Sebba". *Journal of Pidgin and Creole Languages* 11: 123-134.
- Thomason, S. G. and Kaufman, T. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley, CA: University of California Press.
- Thomason, S. G. 1993. "On identifying the sources of creole structures". In *Africanisms in Afro-American language varieties*, S.S. Mufwene (ed.), 280-295. Athens, GA: University of Georgia Press.
- Thompson, S. 1988. "A discourse approach to the cross-linguistic category 'adjective'". In *Explaining linguistic universals*, J.A. Hawkins (ed.), 167-185. Oxford: Basil Blackwell.
- Van Stipriaan, A. 1993. *Surinaams contrast, Roofbouw en overleven in een Caraibische plantagekolonie, 1750-1863*. Leiden: KITLV.
- Veenstra, T. 1996. *Serial verbs in Saramaccan: Predication and creole genesis*. The Hague: HIL.
- Voorhoeve, J. 1957. "The verbal system in Sranan". *Lingua* 6: 374-396.
- Voorhoeve, J. 1962. *Sranan syntax*. Amsterdam: North-Holland.
- Waite, J. 1983. "Verbs, adjectives and predicate marking in Isle de France creole: Back to basics". *Te Reo*, 26, 59-64.
- Winford, D. 1993. *Predication in Caribbean English creoles*. Amsterdam: John Benjamins.
- Winford, D. 1997. "Property items and predication in Sranan". *Journal of Pidgin and Creole Languages* 12: 237-301.
- Winford, D. 1998. "Creoles in the context of contact linguistics". To appear in *Pidgin and creole linguistics in the 21st century*, G. Gilbert (ed.). Amsterdam: John Benjamins.

Variable Concord in Portuguese: The Situation in Brazil and Portugal

Anthony J. Naro and Maria Marta Pereira Scherre

1. Introduction

Non-standard dialects of Brazilian Portuguese throughout the country show certain variable agreement or concord phenomena that are of disputed origin. Examples are variable subject/verb number agreement, variable number agreement within the noun phrase, and variable number agreement in the predicate phrase (Naro 1981; Scherre & Naro 1991). Examples are:

- (1) Variable subject/verb number agreement:

eles ganhaM (3rd pl.) *demais da conta* ‘They get much too much’
eles ganha (3rd sg.) *demais* ‘They get too much’

- (2) Variable number agreement within the noun phrase:

oS (pl.) *freguesES* (pl.) ‘the customers’; *aS* (pl.) *boaS* (pl.) *açÕES* (pl.) ‘the good actions’
aS (pl.) *codorna* (sg.) ‘the game hens’; *aS* (pl.) *porta* (sg.) *aberta* (sg.) ‘the open doors’
essaS (pl.) *estradaS* (pl.) *nova* (sg.) ‘these new roads’; *do* (sg.) *meuS* (pl.) *paiS* (pl.) ‘of (the) my parents’¹

- (3) Variable number agreement in the predicate phrase:

aS (pl.) *coisaS* (pl.) *tÃO* (pl.) *muito caraS* (pl.), *né?* ‘Things are very expensive, aren’t they?’
aS (pl.) *coisa* (sg.) *tÁ* (sg.) *cara* (sg.) ‘Things are expensive’

Recently, variable gender agreement has also been found in certain isolated

rural dialects, which show nearly complete variable leveling of the verbal paradigm (Baxter & Lucchesi 1993; 1997).

2. Drift versus Creolization

There are basically two schools of thought with respect to the origin of these instances of variable lack of concord (or marking), as well as other variable phenomena in non-standard Brazilian Portuguese. One view has it that nothing more is at work in the origin of these variable phenomena than the centuries-old drift of the Indo-European languages in general, and the Romance languages in particular, toward a less inflected grammar. According to this point of view, the variable structures exemplified above for Brazil might have existed in Portugal as well as in Brazil, although perhaps in a different time frame and to a different extent (Naro 1981; Tarallo 1993).

The other view is that Brazilian Portuguese is radically divergent from European Portuguese in this respect. The principal cause of the divergence from European lines of development under this view is the massive presence in Brazil of people of African origin, whose original languages might have influenced Brazilian Portuguese through a hypothetical pidgin or creole Portuguese stage (Guy 1989). The only other major ethnic group that could have participated in a hypothetical pidgin or creole Portuguese would be the Amerindians, but they are known to have utilized a Tupi-based system called *Língua Geral* both to communicate with Europeans and with diverse Amerindian populations.² They had no need for a Portuguese-based system. Indeed, in the early colony nearly all Europeans were fluent in *Língua Geral*.

The creolization approach is difficult to pin down to specifics, because the term has never had a single widely accepted definition in linguistics. The classical use of the concept of creolization, requiring two historically separate stages of evolution, has become controversial. Under this conception, there is first a pidgin stage, in which a system of verbal communication that is no one's native language is used for transmission of information between speakers who have no natural language in common. This stage usually involves a vocabulary that comes predominantly from one particular language, sometimes called the base language. Inherent in this notion of pidgin is the lack of a stable common norm, so that any structure that works for purposes of communication at any moment is acceptable at that time. This sort of grammatical

anarchy has been thought to typically lead to an intermediate phase of relatively stable grammatical norms, even if the pidgin continues to be no one's native tongue. In the classical definition, the second stage of the process, creolization, occurs when the pidgin becomes the native language of a segment of the community. However, given that stabilization of a pidgin can occur without creolization in the classical sense of the creation of a community of native speakers, the notion of creolization loses its linguistic application and becomes a merely external historical notion.

Although the circumstances of pidgin/creole usage are typically thought of as involving simplification (sometimes called 'reduction') of grammatical structure, it is difficult to identify the origin or cause of this phenomenon, or indeed even to state which grammar has undergone simplification, because the base language need supply only vocabulary, not necessarily grammar, during the normless pidgin stage. Given this fact, together with the lack of any clear linguistic content to the classical notion of creolization, an attenuated use of the term has evolved in recent linguistic work (cf. Holm 1992; Baxter & Lucchesi 1997). In this new usage, creolization no longer is identified with nativization of a system formerly not a native language for anyone, but rather with language acquisition in the context of any sort of irregular or discontinuous transmission of language.

We, however, see no gain of insight in this new usage. It seems to us that this attenuated notion of creolization has no empirical content, since nearly all linguistic history, with the possible exception of a village completely isolated from all human interchange, would qualify as creolization to a greater or lesser extent. Therefore, when we use the term creolization in this paper we will be referring to the classical notion.

A conciliatory approach as to the origins of popular Brazilian Portuguese, adopted by Naro & Scherre (1993), postulates that the original source of variable concord phenomena came from Portugal, but that conditions of endemic pidginization and adult second language learning that predominated throughout the history of Brazil, even before the arrival of Africans, accelerated and exaggerated³ the original trend during the process of nativization of the Portuguese language by communities of particularly diverse backgrounds. It should be noted that we are using the term *nativization* strictly in the sense of change from a non-native to native language of a speech community. In particular, our usage of the term excludes any assumptions as to the nature of the language in its non-native forms. We do not require, in our definition, that

the non-native forms that undergo nativization need be reduced or simplified in any consistent manner in the community, as is often assumed in pidgin/creole studies.⁴ Our terminology also excludes the assumption that nativization is an instantaneous process with uniform results throughout the community. However, since we do not have access to data from the intermediate stages of evolution of Brazilian Portuguese between the sixteenth and the twentieth centuries, we will regrettably be unable to take these stages into account in any meaningful way.

3. Variable Agreement in Modern European Portuguese

In Naro & Scherre (1993: 441-444) we reviewed the historical evidence for European origins of loss of agreement marking in Portuguese, tracing developments from pre-Latin times to pre-Classical Portuguese. We presented firm evidence that loss of final *-s* and nasalization can be found in stages of Latin and earlier languages from which Portuguese is derived, as well as in modern European Portuguese itself. We showed that at the present time variable agreement phenomena in Brazil have both phonological and morphological components, and presented evidence that the morphological component is later than, and derived from, the phonological one.

Despite the historically documented evidence we have presented in the printed literature, variable agreement phenomena of the sort under discussion here are generally thought by linguists to be exclusively Brazilian in origin. European Portuguese scholars, even those whose main field of interest is dialectology, have always been nearly unanimous in declaring that all agreement rules are categorical in all dialects of Portugal. The situation continues much the same today — on a recent trip to Portugal (September/October 1995), all of the professional linguists and dialect researchers we met with stated unambiguously that the agreement rules are always used categorically throughout Portugal by all social classes and on all levels of education, including illiterate speakers.

Only two exceptions are generally admitted in the literature and were also acknowledged by all of the Portuguese researchers we contacted: 1) variable use of 3rd singular and 1st plural verb desinences with *a gente*, formally a third person noun, but used as a replacement for the 1st plural pronoun *nós* 'we' in both Portugal and Brazil (*a gente fala* (3rd singular) or *falamos* (1st

plural) ‘we speak’ [Nascimento 1987: 264–266];⁵ 2) variable use of final -s in some southern regions of Portugal near Andalucia, especially in the town of Barrancos, which once belonged to Spain (Saramago 1994).

Despite the monolithic anti-variation front we encountered among our fellow linguists in Portugal, there are some exceptions in the literature. The most notable is a monograph entitled *A linguagem dos pescadores da Ericeira* by Joana Lopes Alves, first published in Lisbon in 1965 (Alves 1993).⁶ Ericeira is located on the Atlantic coast, north of Lisbon, near the well-known tourist destinations of Mafra and Sintra. At the time of Alves’s original research, Ericeira was a quiet fishing village (but has since developed into a tourist attraction). Among other interesting characteristics of the local dialect, Alves notes:

The 3rd person singular is sometimes used instead of the 1st person, as in *eu ontem foi* [3rd sg.] *à Malhada* ‘I went to the shore yesterday’, *eu na quinta-feira apanhou* [3rd sg.] *2 kilos de pólvos* ‘on Thursday I got two kilos of octopus’ (Lopes Alves 1965: 190).

The 3rd person singular is sometimes used instead of the 3rd plural, in examples such as *as quenguerelas só presta* [3rd sg.] *para pescar* ‘the quenguerelas (a type of shell fish) are useful only for fishing’; *lanço das duas pedras que tá* [3rd sg.] *no meio do lanço* ‘I throw from the two stones which are in the middle of the net’; *corre* [3rd sg.] *todos os seus criados / a dar água à Adelina* ‘all of the servants run to give water to Adelina’ (ibid.190).

The use of the singular, when the correct form should be plural, is frequent: *foi há muito* [sg.] *ano* [sg.] ‘it was many years ago’, *andê por munto* [sg.] *sito* [sg.] ‘I was in many places’; *tenho cinquenta ê um ano* [sg.] ‘I am fifty one years old’ (ibid. 191).

Although the original corpus was not recorded on tape, we had the opportunity to examine other recordings from the 1980s briefly together with Alves, to whom we wish to express our gratitude.⁷ In them we found a rich source of variable data. Consider, for example, the following selections from a continuous stretch of the speech of a fisherman who was 69 years old in 1982 and had some elementary schooling in his youth. Note in particular the variable marking of the verb *ser* ‘to be’ with plural predicates — in (4) with imperfect tense forms and in (5) with present tense forms. The noun phrases *três sardinha* ‘three sardines’ and *os pixeru* ‘the fish-mongers’ in (4) are both used without the plural -s on the head in typical Brazilian style, along lines noted by Alves above. We have also included some noun phrases that are fully marked for plural in order to exemplify the variation to be found in the

recordings:

- (4) *Era* (3rd sg.) *duas, três sardinha* ‘There were two or three sardines’
Éramos (1st pl.) *três, quatro homis, né?* ‘We were three or four men, right?’
Eru (3rd pl.) *os pixeru logo* ‘It was the fish-mongers quickly’
- (5) *São* (3rd pl.) *cinco camaradas* ‘they are five partners’
É (3rd sg.) *duas partes pru dono e uma pa cada um da gente* ‘There are two parts for the owner and one for each of us’

Another speaker from Ericeira, a woman who was 84 in 1982, uttered the following clause, in which a third singular salient verb form (*faltou/faltaram* ‘missing’) is used with a plural coordinate subject:

- (6) *Mas, minha senhora, pescadas e linguado nunca lá faltô* (sg.) *a eles*
 ‘But, ma’am, they were never wanting for hake or sole’.

We also listened to tape recordings from other parts of Portugal that were kindly made available to us by researchers from the *Atlas Lingüístico-Etnográfico de Portugal e da Galiza* (ALEPG) and *Português Fundamental* (Nascimento, Marques & Cruz 1985) projects. Here too we found variable data. In (7) below we list some further examples of the same type as *os pixeru* ‘the fish-mongers’, that is, plural noun phrases without the plural marker *-s* from the head onwards.⁸

- (7) a. Speaker from the District of Guarda (Nave de Haver), north of Portugal:
nabiças pequena ‘small turnips’
os vivo ‘living people’
- b. Speaker from Alenquer, near Lisbon:
três ou quatro mil quilo ‘three or four thousand kilos’
- c. Speaker from Pavia, center-south of Portugal:
umas pouca ‘a few’
uns limõe ‘some lemons’

Examination of the literature on Portuguese dialects reveals further instances of non-agreeing forms throughout the country. From Odeleite, in the south, we find *condo morria* (3rd sg.) *pessoas de família chigada* ‘when people from the close family died’ (Cruz 1991: 159); from Azóia, near Sintra,

as borricêras que viero (3rd pl.) *onte é que fez* (3rd sg.) *isto* ‘the drizzles that came yesterday did this’ (Marques 1968: 61).

In connection with the use of partially unmarked plural noun phrases, it is interesting to note that Mira Mateus registered the form *as raízes* (fem. pl.) *enterrado* (masc. sg.) *na carne* ‘the roots buried in the flesh’ in her unpublished 1954 dissertation on the vernacular usage of Lisbon (Mira 1954: 150). In this example the adjective *enterrado* ‘buried’ agrees in neither gender nor number with its head *raízes* ‘roots’. She also cites the clause *os nossos agasalhos é* (3rd sg.) *estes* ‘our outer garments are these’, in which the verb does not agree with the subject. In fact, Mira Mateus states that “in popular speech, lack of agreement, considered an error from the grammatical point of view, is frequent”. We find the use of the term ‘frequent’ (parallel to Alves & Cruz, cited earlier) in Mira Mateus’s comment quite striking, given that colleagues we met on our trip to Portugal, in personal communication, were so vehement in rejecting the possibility of the use of non-agreeing forms in Portugal under any circumstances.

As mentioned above, variation in agreement has both morphological and phonological components in Brazil. We find evidence for the same characteristic in Portugal. On the one hand, forms such as *faltou* (3rd sg.) in place of *faltaram* (3rd pl.) ‘were wanting’ in (5) or *é* ‘is’ in place of *são* ‘are’ in (6) could only be viewed as substitution on the morphological level. On the other hand, in the form *uns limõe* ‘some lemons’ in (7), the singular ending *-ão* in *limão* ‘lemon’ has yielded to the plural ending *-ões*, as in *limões* ‘lemons’, but the final *-s* is missing.

It is our impression that variable agreement can be found throughout the entire territory of Portugal, both within the noun phrase and between the verb and the subject. The reasons for the scarcity of references to this phenomenon in the literature, and our Portuguese colleagues’ intuitions’ denial of its existence, are unclear. The non-agreeing forms are statistically rare, but we encountered no difficulty locating and checking a reasonable number of them in only a few days of listening to tape recordings (in the company of some of the same scholars who had previously denied their existence). It is clear that much work remains to be done in order to determine the geographical and structural dimensions of variable concord agreement in present-day Portugal, but its existence can no longer be denied. Variation in agreement is not an exclusively Brazilian phenomenon in the synchronic grammar of modern Portuguese.

4. Historical Documentation of Variable Agreement in European Portuguese

Texts from earlier stages of Portuguese, particularly those from the period before grammatical norms were established during the classical stage of the sixteenth century, also occasionally show lack of agreement markers (Mattos e Silva 1991). These instances are normally viewed as scribal or printing errors by modern scholars, who almost without exception feel it is their duty to edit the texts for publication, changing the original occurrences of singular verb forms to plural forms in plural contexts, as the modern and classical norms prescribe.

In a study of eight pre-classical texts (Anonymous 1515, Aluarez 1540, Cintra 1954, Magne 1955, Lopes 1968, Mira Mateus 1970, Mattos e Silva 1971), we have found over two hundred occurrences of third singular verb forms in environments where the norm would categorically require third plural forms. The proportion of non-agreeing data varies somewhat from text to text, but is usually less than 1%. In the *Demanda do Santo Graal* (the Portuguese version of the Holy Grail legend), for example, there are about 20 non-agreeing tokens and 4700 agreeing tokens, so that the rate of lack of agreement in this text is approximately 0.4%.

In (8) - (11), we cite four examples taken from ancient Portuguese texts. The first two are types that are frequent in our diachronic sources; the second two are rarer:

- (8) *...mas DAVA* (3rd sg.) *lhe gram torva as portas çarradas*
 'but the closed gates caused him much consternation' (Mira Mateus 1970: 244, cxxxix)
- (9) *...a todos aqueles que se fazem* (3rd pl.) *afora da carreira do pecado e TORNA* (3rd sg.) *a direita carreira*
 'to all those who abandon the road of sin and turn to the right road' (Magne 1955: xxi, 160, 223)
- (10) *Entom os parentes OUVE* (3rd sg.) *conselho e confessaron* (3rd pl.)
 'Then the relatives conferred and confessed' (Mattos e Silva 1971: 35)
- (11) *aqueles que, da gente d'Alexandria, REINO* (3rd sg.) *no Egipto*
 'those who, among the people of Alexandria, reigned in Egypt'.
 (Mira Mateus 1970: 554)

The verb form *dava* ‘gave’ in (8) would need only the addition of a final nasal to make it formally a plural. It occurs with a post-posed inanimate subject. The verb *torna* ‘turn’ in (9) similarly is missing only the final nasal, but has a pre-posed subject that is separated from it by many syllables of intervening phonic material. In both cases, the structural circumstances would highly favor lack of marking in modern Brazilian Portuguese. In (10), on the other hand, we find a verb form *houve* ‘had’ that is very different from its plural form *houveram*. It occurs with a directly pre-posed human subject. Finally, in (11) we find another verb form *reinou* ‘reigned’ that is quite different from the corresponding plural *reinaram*, but with a distant human subject. In modern Brazil, the last two contexts would not favor lack of agreement, although it would certainly be possible. Correspondingly, such forms are rare in the ancient texts.

The general impression one gets upon inspecting the 235 tokens of lack of agreement we have collected from medieval Portugal is that they could still be used naturally today in more or less the same relative proportions, at least in Brazil, provided the vocabulary were modernized. To test this intuition in a more rigorous way, we coded each non-agreeing token in accord with categories known to be statistically valid today in Brazil. We then took a random sample containing 639 agreeing plural tokens in the same texts, and ran the pool of data obtained in this way through the GOLDVARB/VARBRUL (cf. Pintzuk 1988, Sankoff 1988) variable rule program. Naturally, the resulting raw frequencies in terms of percentage of plural use grossly overestimate the frequency of non-agreeing usage. However, the calculated relative weights (with the exception of the ‘input’) are correct, and would not change significantly with the addition of more agreeing plural data (see Sankoff 1988 on the statistical treatment of rare variants). Given this situation, in the tables for data from the medieval texts we report only the number of non-agreeing tokens we found and the corresponding relative weights.

One effect we found to be statistically significant in the medieval written data may be familiar from the published literature on spoken Brazilian Portuguese: phonic saliency of the singular/plural opposition. In Brazil, more frequent use of singular forms in plural contexts occurs when this opposition is realized by a less perceptible difference in form between the singular and plural form (e.g. *dava/davam* ‘he gave’/‘they gave’; *torna/tornam* ‘he turns/they turn’). This basic observation led to the postulation of a detailed analysis of the saliency hierarchy, containing six degrees, or levels, of the phonic

opposition between the singular and plural forms of the verb. The detailed version of the saliency hierarchy has been shown to be valid in the speech of both literate and illiterate speakers in Rio de Janeiro as well as in many other localities throughout Brazil. Nonetheless, the strongest dimension of the hierarchy — stress of the desinence — can be used to reduce it to only two levels, as outlined in Naro (1981: 74): “the first level contains those pairs in which the phonetic segments that realize the opposition are unstressed in both members” (less salient opposition: *come/comem* ‘he eats/they eat’; *ganha/ganham* ‘he wins/they win’; *falava/falavam* ‘he spoke, they spoke’; *faz/fazem* ‘he does/they do’.); “the second level contains pairs in which these segments are stressed in at least one member of the opposition” (more salient opposition: *dá/dão* ‘he gives/they give’; *comeu/comeram* ‘he ate/they ate’; *ganhou/ganharam* ‘he won/they won’; *é/são* ‘he is/they are’; *disse/disseram* ‘he said/they said’).

Table 1 shows our results for ancient Portuguese texts. In this table the column labeled N reports the number of occurrences of non-agreeing forms we found for each independent variable; the ‘Relative weight’ column reports the relative weights calculated and chosen as statistically significant by GOLDVARB/VARBRUL:

Table 1. Rate of subject/verb concord according to saliency of the singular/plural opposition in eight medieval Portuguese texts

Category	N	Rel. weight
–salient	204	0.33
+salient	31	0.75

For purposes of comparison, in Table 2 we present the corresponding results for the speech of illiterate and literate speakers in present-day Rio de Janeiro:

Table 2. Rate of subject/verb concord according to saliency of the singular/plural opposition for illiterate (Naro 1981) and literate speakers in Rio de Janeiro (Naro & Scherre 1996)

Category	Illiterate speakers		Literate speakers	
	Frequency	Weight	Frequency	Weight
–salient	972/3568=27%	0.22	1549/2496=62%	0.32
+salient	2030/2742=74%	0.78	1820/2136=85%	0.71

The results in Tables 1 and 2 demonstrate that in ancient Portuguese texts, just as in modern spoken Brazilian Portuguese, plural marking was used less frequently with low saliency verb forms. Even though the frequencies may vary widely in the three cases, the relative weights are essentially equal, and show that the factors that controlled the variation were the same in medieval Portugal as they are today in modern Brazil.

Another factor group that may be familiar is position of the subject with respect to the verb, with more frequent use of singular verb forms in the context of plural subjects when the subjects are distant or post-posed to the verb. We coded the following groupings for distance of pre-posed subjects with respect to the verb: 1) zero to two syllables of intervening phonic material, 2) three to ten syllables, 3) more than ten syllables. Our results for this variable are shown in Table 3:

Table 3. Rate of subject/verb concord according to position of the subject with respect to the verb in eight medieval Portuguese texts

<i>Category</i>	<i>N</i>	<i>Rel. weight</i>
0-2 syllables	27	0.71
3-10 syllables	33	0.46
>10 syllables	10	0.19
post-posed	42	0.43
zero ⁹	123	0.36

In Table 4 we present the corresponding results for subject position in the spoken language of modern Brazil. Unfortunately, in the case of this variable, the continuum of distance between the subject and the verb has been categorized in slightly different ways in each research project, but the essentially identical functioning of distance is striking:

Table 4. Rate of subject/verb concord according to position of the subject with respect to the verb for illiterate (Naro 1981) and literate speakers in Rio de Janeiro (Naro & Scherre 1996)

<i>Category</i>	<i>Frequency</i>	<i>Rel. weight</i>
<i>Illiterate speakers</i>		
0-5 syl	1802/3680=49%	0.71
>5 syl	84/ 254=33%	0.41
post-posed	59/ 252=23%	0.24
zero	1057/2124=50%	0.65

<i>Category</i>	<i>Frequency</i>	<i>Rel. weight</i>
<i>Literate speakers</i>		
0 syl	1529/1857=82%	0.61
1-5 syl	778/1064=73%	0.54
>5 syl	61/ 97=63%	0.40
post-posed	50/ 194=26%	0.08
zero-far	220/ 255=86%	0.67
zero-close	731/1166=63%	0.36

In all three instances presented in Tables 3 and 4, it is clear that there is less use of agreeing forms as the subject becomes more distant from the verb or is post-posed to it. Nonetheless, the effect of the post-posed category is not as strong in the ancient texts as in the modern spoken language (an effect previously noted by Mattos e Silva [1991:71] for a different type of construction).

Other variables that are equally valid for both modern Brazilian and ancient Portuguese may be less familiar because their importance in Brazilian Portuguese has only recently been discovered (Scherre & Naro 1996; 1998). One such category is the feature [human] of the subject, with much more use of plural marking for human plural subjects. Results for the medieval texts are shown in Table 5:

Table 5. Rate of subject/verb concord according to the feature 'human' of the subject in eight medieval Portuguese texts

<i>Category</i>	<i>N</i>	<i>Rel. weight</i>
human	148	0.58
non-human	87	0.27

Table 6 exhibits the corresponding results for literate speakers in Rio de Janeiro.¹⁰

Table 6. Rate of subject/verb concord according to the feature 'human' of the subject for literate speakers in Rio de Janeiro (Scherre & Naro 1998)

<i>Category</i>	<i>N</i>	<i>Rel. weight</i>
human	3017/3981=76%	0.53
non-human	264/ 505=52%	0.29

As Table 5 and 6 show, in the medieval Portuguese texts, just as in the modern spoken language of Brazil, plural human subjects are far more likely than non-human subjects to impose an explicit plural agreement marker on the verb.

It should be stated at this point that we have not been able to test all of the independent factors known to be valid in modern spoken Brazilian Portuguese on the two hundred or so tokens we have found in medieval Portuguese documents. For example, we simply do not find the type of topic chain data that would permit us to test the parallel processing effects that are so strong in Brazilian Portuguese, where marked forms normally lead to further marked forms, and unmarked forms to more unmarked forms (Scherre & Naro 1991). In the texts we find only a very few topic chains, most containing only two clauses, and in these there are more with the second verb unmarked (see example [9]) than with the first verb unmarked (see example [10]).

Another effect we have been unable to verify in the ancient texts is the strong tendency for the relativizer *que* 'that' to decrease chances of marking when it intervenes between the explicit subject in the main clause and the dependent verb, as in (9) above (repeated here as [12] with the relativizer and the dependent verb in small capitals):

- (12) *a todos aqueles QUE se FAZEM* (3rd pl.) *afora da carreira do pecado*
'to all those who abandon the road of sin'

Although this effect seems to exist in the medieval data, both in terms of number of tokens and relative weights, statistical significance at the 0.05 level is not attained with the data presently available.

There are some effects in the medieval Portuguese data that seem stronger than in Brazil, such as the tendency for forms of *ser* 'to be' to agree with the predicate rather than the subject. An example is:

- (13) *e as demais das ervas destas mōtanhas e ribeiras HE mangeriçō*
'and the remaining herbs of these mountains and rivers is marjoram' (Alvarez: fl.4b, c.7)

Finally, there are areas where possible differences between the modern and medieval varieties remain to be tested. In the ancient texts, intransitive and sentential complement verbs very rarely lack explicit plural markers in the appropriate environments. There is also a statistically significant effect that favors plural marked verbs with plural objects, given that the subject is plural.

We do not know if these effects are valid in Brazil as we have not yet studied them. In any case, it must be recalled that, even within contemporary Portugal and Brazil, the written and spoken languages of each country are very different from each other. If we find some differences between the two varieties being studied here, we must bear in mind that we are dealing with distinct registers as well as different locations in time and space.

5. Implications

Returning to our point of departure in this paper, the view of variable agreement in Brazil as being a phenomenon of a radically non-European type is falsified by the results presented here, which show that variable agreement in both present-day spoken European Portuguese and medieval written European Portuguese exhibits the most fundamental structural characteristics found in modern spoken Brazilian Portuguese. The differences between the European and Brazilian phenomena are a matter of degree, not type. The general diachronic picture is preservation of the hierarchical effect of the conditioning factors, accompanied by an increase over time in Brazil of the overall average level of use of the unmarked variant of the verb in the context of a plural subject. No radical, or even light, restructuring has occurred in Brazil since the types of structures used have not changed; the change was basically in the frequency of use of the available types.

In terms of overall frequency, this situation is the inverse of that found by Oliveira e Silva (1982; 1996) for the historical development of the use of the article before possessives in phrases of the type *o meu filho Artur* '(the) my son Arthur': from the fifteenth to the twentieth century, the use of the article in this context increased continually in the written language of Portugal, starting out at about 10% and reaching approximately 95% at the present time. During the entire course of this evolution the conditioning factors did not change in Portugal, remaining throughout the same as those still valid today in Brazil. However, in Brazil, the general level of use of articles before possessives remained essentially stable, with the result that they are now used less in Brazil than in Portugal in this environment.

Since the evolution of Portuguese in Brazil is often compared to creolization, in both the classical and attenuated senses mentioned at the outset of this paper, we have attempted to identify a clear creolization situation that might

be similar to that of Brazil. In this connection, the dynamics of Portuguese language nativization in Brazil evidently produced the same sort of results as did Tok Pisin nativization in New Guinea, even though the two situations are vastly different in nature and, in particular, do not share the same type of pre-nativization circumstances. Sankoff (1980) has shown that in New Guinea the nativization of the pidgin by a new generation of children born into households with parents who did not have Tok Pisin as their native language did not cause any restructuring or other radical changes in the children's grammar with respect to that of their parents. Rather, the new generation of native speakers "appear to be carrying further tendencies which were already present in the language" (Sankoff & Laberge 1980: 208). According to traditional terminology, the adults would be classified as "pidgin" speakers and their children as "creole" speakers. At the time of Sankoff's research, "creolization", i.e., nativization of the pidgin was underway.

In the sense of "carrying further tendencies which were already present" — and in this sense only — we see a parallel between the evolution of Portuguese in Brazil and the process of creolization in New Guinea. However, history does not record the existence in Brazil of any form of Portuguese, creolized or not, associated exclusively with speakers of African or any other origin. While it is true that there are isolated rural communities of black people in the state of Bahia who use very little concord in their speech (Baxter & Lucchesi 1993, 1997), the tendency toward infrequent usage of concord can be found throughout all rural areas of Brazil (Amaral 1920; Veado 1982), from north to south, whether predominantly European, African, Asian, Amerindian, or mixed in origin. In particular, there are isolated rural communities of white people in Bahia, such as Mato Grosso (municipality of Rio de Contas), which show the same sort of structures typically associated with creolization (variable use of all types of agreement, including gender, use of a second negative *não* 'not' at the end of a negative sentence¹¹, paradigm reduction in verbs, the so-called all purpose preposition *ni*, and so on (see Callou 1998).¹²)

The use of the term "creolization" in Brazil would be misleading, since there can be no association of the process with any particular substrate group or groups who might have influenced the evolving language in a consistent way. Furthermore, there is no evidence to suggest the existence of a previous Portuguese-based pidgin that could have contributed structural peculiarities. The term "nativization", in the sense in which we are using it, i.e., to mean

“change from a non-native to a native language of a speech community”, accurately describes what actually happened to the Portuguese language in Brazil. Given these facts, we explicitly reject the position that Brazilian Portuguese is a semi-creole, has a creole history, or has undergone light creolization (Silva Neto 1986, Câmara 1975, Jeroslow 1975, Guy 1989, Holm 1992, Baxter & Lucchesi 1993, Ferreira 1994).

We wish to emphasize once again that in the case of Brazil, all of the features present at the current stage of the historical process of evolution of variable concordance were also present at its beginning. The basic change was in the general overall tendency in the frequency of use of the non-agreeing forms. In more technical terms, the variable constraints that govern the use of agreement did not change over time; only the input weight changed. In this respect, the process is qualitatively identical to normal language change (Kroch 1989).

It is perhaps worth mentioning at this point that, aside from variable concordance, many other so-called peculiarities of Brazilian Portuguese can be easily traced back to Portugal, especially to the non-standard varieties documented in studies of linguistic geography. This is true, in particular, of use of nominative pronouns as accusatives (*ele* ‘he’ used instead of *o* ‘him’ as direct object [Marques 1968: 56; Alves 1993: 180; 189]); use of oblique pronouns as nominatives (*mim* ‘me’ instead of *eu* ‘I’ as subject [Cruz 1991: 153]); use of the third person reflexive *se* as a universal reflexive pronoun at the expense of first and second person reflexive forms (Marques 1968: 56); *ter* ‘to have’ used as an existential (Marques 1968: 58); expanded use of periphrastic verb forms such as the future tense with *ir* ‘to go’ (Cruz 1991: 168; Marques 1968: 58); general reduction of tense and mood forms (Alves 1993: 183, 190); greater frequency of use of coordination and juxtaposition, with less use of subordination, in syntax (Mira 1954: 144-145; Delgado 1970: 158); frequent use of expletives and other emphatic processes (Mira 1954: 165-170; Marques 1968: 58); and preferential use of prepositional forms related to Latin *in* and other changes in the use of prepositions (Marques 1968: 60; Cruz 1991: 176-177). We have not yet identified any feature of Brazilian Portuguese that has no clear ancestor in Portugal.

In the medieval written data analyzed in this paper, non-agreeing forms constitute only about 1% of all attested usage. Unfortunately, we have no way of determining what the corresponding rate in the spoken language may have been at the time. Nonetheless, when overall variation is very small, from the

point of view of certain theoretical perspectives it might be tempting to postulate that the grammatical mechanism of concordance is categorical and that non-agreeing tokens arise through performance interference with competence. The features that control variation in the use of agreement — phonic salience, distance between subject and verb, and the feature ‘human’ of the subject¹³ — are stated in such a way as to be universally applicable and independent of any particular structural feature of the Portuguese language. For this reason, the postulated performance mechanism that interferes with competence might be universal. A problem arises, however, when the input parameter reaches levels above 50%, as in many varieties of modern Brazilian Portuguese. Is there any insight to be gained from classifying the variation as due to performance, rather than competence, at this stage? If not, at what input level is there a switch from performance to competence? Given the current state of the art in linguistic theory, we see little empirical content to these ideas, but if they are adopted, the fact that so-called performance is regular and rule-governed must not be lost from view.

6. Conclusion

At the outset of this paper we mentioned two schools of thought, apparently in conflict, as to the nature of the origins of non-standard Brazilian Portuguese — on the one hand, exclusively European origin; on the other, pidginization/creolization due to the presence of speakers of African or other non-European languages. The overall pattern of development sketched above is more consistent with European origins, but also depends upon the catalyzing effect of nativization pointed out by Sankoff. Our conclusion is that Modern Brazilian Portuguese is the natural result of the centuries-old drift inherent in the language brought from Portugal, exaggerated, to be sure, in Brazil by extensive contact between adult language speakers of particularly diverse origins and the nativization of Portuguese by communities of such speakers.

Notes

We wish to express our gratitude to our research assistants Ricardo Joseh Lima, Viviane Conceição Antunes, and Elane Barreto dos Santos, without whose work we would not have been able to carry out the massive search through medieval Portuguese texts reported here. We also

wish to thank our Portuguese friends and colleagues for the warm welcome we received during our brief visit to Lisbon in 1995. Our thanks first of all to Maria Helena Mira Mateus for all of her assistance, both personal and academic, and to Joana Alves for the time she took to work with us. Researchers at the Centro de Linguística da Universidade de Lisboa provided us with personal assistance and the use of their office space and other resources. For their exceptional kindness and academic assistance we wish to sincerely thank Maria Fernanda Bacelar do Nascimento, João Saramago, and Maria Luísa Segura da Cruz, as well as the Center's librarians Teresa Loyolla and Adélia.

1. In Portuguese the definite article can optionally be used with possessive phrases. The form *do* is bimorphemic, consisting of the preposition *de* 'of' and the definite article *o* 'the' (masc. sg.)
2. See Naro & Scherre (1993) and, for discussion of facts and polemics, Rodrigues, J. 1983; Rodrigues, A. 1994:99-109.
3. We are here using the terms 'accelerated' and 'exaggerated' in an informal sense, making an implicit comparison between the present-day stage of development of the phenomena under examination in the Portuguese language in Brazil with the present-day state of affairs in Portugal. In absolute terms, we do not know when, in the nearly 500-year history of Brazil, nor with what speed, the changes that differentiate these two sister varieties actually occurred.
4. Silva Neto (1986: 594-595), while discussing this matter, speaks of the 'precipitation' of changes caused by non-native speakers of Portuguese in Brazil and assumes the existence of a creole or semi-creole as a trigger for the process of simplification (cf. 593, 595). Câmara (1975:77) also assumes the existence of a creole, associated with black Africans, attributing to it a similar role as a trigger. However, he states explicitly that the changes follow "the Portuguese language's own structural tendencies". We also postulate that initial tendencies were exaggerated, but our view of the overall process is different than that of the authors cited in that we do not see any evidence of a creole stage.
5. The use of the 1st plural desinence *-mos* with *a gente* is undoubtedly more frequent in Portugal than in Brazil.
6. This monograph is a result of one of the many similar research projects carried out throughout Portugal in the 60's by students of Prof. Lindley Cintra at the Universidade de Lisboa.
7. Joana Lopes Alves, an enthusiastic researcher of the fishing community of Ericeira, was kind enough to meet with us on two occasions in order to discuss her work and listen to tapes recorded in 1982. She granted us permission to use her unpublished data in our paper. Other scholars who also helped us with taped data were Maria Luísa Segura da Cruz and João Saramago. We are extremely grateful for their assistance.
8. See comments on the form *uns limõe* below.
9. The 'zero' categories in these tables are not comparable because they were coded in different ways. For some discussion of the problems involved see Scherre & Naro (1991:25-27).
10. The relevance of this variable had not been recognized at the time the research on illiterate speakers was carried out.

11. This phenomenon is known as 'double negative', although there are other double negative constructions in the standard languages of both Portugal and Brazil.
12. Callou (1998 [see also Isensee 1964]), following Harris (1956), emphasizes the isolated nature of the white community of Mato Grosso. Naturally, there are also black and mixed communities in the area, but the nearest such community is located in a valley more than 1500 meters below. All of the historical documentation discovered to date suggests that the two communities had very little contact over the centuries. In particular, neither community provided labor for the other.
13. These features are all interpretable as dimensions of saliency (cf. Naro & Scherre 1999).

References

- Alvarez, F. 1540. *O Preste Joam das Indias*. Luis Rodrigues.
- Alves, J.L. 1993. *A linguagem dos pescadores da Ericeira*. Lisboa: Junta Distrital de Lisboa: Ramos, Afonso & Moita..
- Amaral, A. 1920. *O dialeto caipira*. São Paulo: O Livro.
- Anonymous. 1515. *Boosco Deleytoso: com privilegio del Rey Nosso Señor*. Lisbon: Hernam Campos.
- Baxter, A. N. and Lucchesi, D. 1993. "Processos de descrioulização no sistema verbal de um dialeto rural brasileiro". *Papia* 2:59-71.
- Baxter, A. N. and Lucchesi, D. 1997. "A relevância dos processos de pidginização e criouliização na formação da língua portuguesa no Brasil". In *Estudos lingüísticos e literários*, 65-84. Salvador: Universidade Federal da Bahia Program in Letters and Linguistics.
- Callou, D.M.I. 1998. "Um estudo em tempo real em dialeto rural brasileiro: questões morfossintáticas". In "*Substandard*" e *mudança no português do Brasil*, Gro?, S. and Zimmermann, K. (eds.), 255-272. Frankfurt am Main: Teo Ferrer de Mesquita.
- Câmara Jr., J.M. 1975. *Dispersos*. Rio de Janeiro: Fundação Getálio Vargas.
- Cintra, L.F.L. 1954. *Cronica geral de Espanha de 1344: Edição crítica*. Lisbon: Academia Portuguesa da História.
- Cruz, M.L.S. da. 1991. *O falar de Odeleite*. Lisboa: Instituto Nacional de Investigação, Centro de Lingüística da Universidade de Lisboa, Barbosa & Xavier.
- Delgado, M.C.S. 1970. *O falar de Baleizão*. University of Lisbon PhD Dissertation.
- Ferreira, C. da S. 1994. "Remanescentes de um falar crioulo brasileiro". In *Diversidade do português do Brasil: Estudos da dialetologia rural e outros*, 21-32. Salvador: Centro Editorial e Didático da UFBA.
- Ferreira, J. de A. 1980. *Alphonse X, Primeira Partida: Édition et étude*. Braga: Instituto Nacional de Investigação Científica.
- Guy, G.R. 1989. "On the nature and origins of popular Brazilian Portuguese". In *Estudios sobre Español de América y Lingüística Afroamericana*, (eds.), 226-244. Bogotá: Instituto Caro y Cuervo.
- Harris, M. 1956. *Town and country in Brazil*. New York: Columbia University Press.
- Holm, J. 1992. "Vernacular Brazilian Portuguese: a semi-creole". In *Actas do Colóquio*

- sobre *Crioulos de Base Lexical Portuguesa*, D'Andrade, E. and Kihm, A. (eds.), 37-6. Lisboa: Colibri.
- Isensee, D.M. 1964. "O falar de Mato Grosso (Bahia): fonêmica - aspectos da morfo-sintaxe e do léxico". Brasília: Universidade de Brasília. Masters thesis.
- Jeroslow, E.H.M. 1975. "Creole characteristics in rural Brazilian Portuguese". Paper presented at the International Conference on Pidgins and Creoles, Honolulu, Hawaii.
- Kroch, A.S. 1989. "Reflexes of grammar in patterns of language change". *Language Variation and Change* 1:199-244.
- Lopes, F. 1968. *Cronica del Rei Dom João I*. Vol. I: Arquivo histórico português, 1915. Vol. II: William J. Entwistle (ed.). Lisbon: Imprensa Nacional.
- Magne S.J., Augusto. 1955. *A demanda do Santo Graal: Reprodução facsimilar e transcrição crítica*. Rio de Janeiro: Instituto Nacional do Livro.
- Marques, M.C.A. 1968. *O falar da Azóia*. University of Lisbon PhD Dissertation.
- Mattos e Silva, R.V. 1971. *Quatro livros dos diálogos de São Gregório: Leitura crítica*. University of São Paulo PhD Dissertation.
- Mattos e Silva, R.V. 1991. "Caminhos de mudança sintático-semântica no português arcaico". *Cadernos de Estudos Linguísticos* 20:59-74.
- Mira, M.H. F. da G. 1954. *Algumas contribuições para um estudo da fonética, morfologia, sintaxe e léxico da linguagem popular de Lisboa*. University of Lisbon PhD Dissertation.
- Mira Mateus, M.H. 1970. *Vida e feitos de Julio César: Edição crítica*. Lisbon: Fundação Calouste Gulbenkian.
- Naro, A.J. 1981. "The social and structural dimensions of a syntactic change". *Language* 57:63-98.
- Naro, A.J. and Scherre, M.M.P. 1993. "Sobre as origens do português popular do Brasil". *Revista DELTA* 9:437-454.
- Naro, A.J. and Scherre, M.M.P. 1996. "Disfluencies in the analysis of speech data". *Language Variation and Change* 8:1-12.
- Naro, A.J. and Scherre, M.M.P. 1999. "Sobre o efeito do princípio da saliência na concordância verbal na fala moderna, na escrita antiga e na escrita moderna". Paper presented at Encontro Nacional de Língua Falada e Escrita III, Universidade Federal de Alagoas, Maceió.
- Nascimento, M.F.B. 1987. *Contribuição para um dicionário de verbos do português: Novas perspectivas metodológicas*. Unpublished dissertation in Portuguese Linguistics presented in fulfillment of requirements for the position of Research Associate. Lisboa: INIC-CLUL.
- Nascimento, M.F.B., Marques, M.L.G. and Cruz, M.L.S. 1985. *Português fundamental*. Lisboa: INIC-CLUL.
- Oliveira e Silva, G.M. de. 1982. *Estudo da regularidade na variação dos possessivos no português do Rio de Janeiro*. Federal University of Rio de Janeiro PhD Dissertation.
- Oliveira e Silva, G.M. de. 1996. "Realização facultativa do artigo definido diante de possessivos e patronímicos". In *Padrões sociolinguísticos: Análises de fenómenos variáveis do português falado no Rio de Janeiro*, In Oliveira, G.M. de and Scherre, M.M.P. (eds.), 119-45. Rio de Janeiro: Tempo Brasileiro.
- Pintzuk, S. 1988. "VARBRUL programs". MS.

- Rodrigues, A.D. 1994. *Línguas brasileiras*. São Paulo: Loyola.
- Rodrigues, J.H. 1983. "A vitória da língua portuguesa no Brasil colonial". *Humanidades* 1: 21-41.
- Saramago, J. 1994. "O atlas lingüístico-etnográfico de Portugal e da Galiza (ALEPG)". In *Geolinguística: Trabajos europeos*, Mouton, P.G. (ed), 225-232. Madrid: Consejo Superior de Investigaciones Científicas.
- Sankoff, D. 1988. "Variable rules". In *Sociolinguistics: An International Handbook of the Science of Language and Society*, Ammon, U., Dittmar, N. and Mattheier, K.J. (eds.), 984-998. Berlin: Walter de Gruyter.
- Sankoff, G. and Laberge, S. 1980. "The acquisition of native speakers by a language". In *The social life of language*, G. Sankoff (ed), 195-209. Philadelphia: University of Pennsylvania Press.
- Scherre, M.M.P. and Naro, A.J. 1991. "Marking in discourse: birds of a feather". *Language Variation and Change* 3:23-32.
- Scherre, M.M.P. and Naro, A.J. 1996. "On the role of the feature [+–] human in variable concord". Paper presented at the New Ways of Analyzing Variation conference, Las Vegas, Nevada.
- Scherre, M.M.P. and Naro, A.J. 1997. "Restrições sintáticas e semânticas na determinação do controle da concordância verbal de número em português". *Fórum lingüístico* 1, 45-71. Florianópolis, Brazil: Federal University Press.
- Silva Neto, S. da. 1986. *História da língua portuguesa*. Rio de Janeiro: Presença.
- Tarallo, F. 1983. "Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias". In *Português brasileiro: Uma viagem diacrónica*, Roberts, I. and Kato, M.A. (eds.), 35-68. São Paulo: Editora da Unicamp.
- Veado, R.M.A. 1982. *Comportamento lingüístico do dialeto rural*. Belo Horizonte: UFMG/PROED.

Nativization and the genesis of Hawaiian Creole

Sarah Julianne Roberts¹

1. Introduction

This article gives a preliminary account of the genesis of Hawai‘i Creole English (HCE) in the early decades of the 20th century. I argue that nativization (the process by which a language acquires a native-speaking community) played an important role in the structural elaboration which distinguished HCE from the pidgin of foreign-born adults. This claim is supported by linguistic evidence, demographic data, and sociological remarks by observers and native-born speakers. I also show that the classic Bickertonian conception of nativization, in which children of immigrants abruptly acquire their parents’ pidgin as their mother tongue, is inconsistent with observed facts. Instead I propose that nativization followed the pattern displayed by a number of other creoles and unfolded in stages, crucially involving the grandchildren of immigrants.

Before I delve into the relevant data, I will briefly raise some theoretical and terminological issues regarding nativization and creole formation which have remained controversial in creolistics.²

2. Creoles and Nativization

Early work on creoles defined these languages as nativized pidgins (Hall 1966; Hymes 1971:3). A growing number of theorists have rejected this definition, as well as the notion that creoles are typologically distinct from pidgins. It is true that all languages classified as creoles have native speakers. But the differences between native and non-native Solomons Pijin, Tok Pisin,

and Sango (languages which have recently acquired native-speaking communities) are mostly limited to tempo and morphophonemics (Sankoff 1977, Samarin 1993, Jourdan & Keesing 1997). These structural differences are usually not regarded as a sufficient basis for the classification of pidgins and creoles as separate classes of language.

It is clear in the abovementioned cases that the major structural innovations setting them apart from early-stage pidgins developed prior to the emergence of native-speaking communities. Jourdan & Keesing (1997:404) note that Solomons Pijin, as spoken by nonnative adults in the 1930s, “incorporated much of the grammatical richness of indigenous languages — in the global syntactic possibilities, the range of complex multi-clausal constructional patterns, and the grammaticalized marking of transitivity and causatives,” as well as tense and aspect auxiliaries. While later native-born speakers may have regularized these patterns, they did not play a part in their innovation.

Historical research on older creoles (including Sranan, Mauritian, and Haitian) suggests that they developed in a manner similar to those undergoing nativization today. Baker (1995:17), after comparing the developmental rates of seven pidgins and creoles, found that nativization “had a negligible effect on speed of development.” Numerous other studies have found that creole formation was a gradual process which spanned between nonnative and native generations (Carden & Stewart 1988; Arends 1989; Plag 1995). Singler (1986, 1990, 1992, 1995) described the nativization process as slow for certain older Atlantic creoles. On this view locally-born speakers were bilingual in both the pidgin-creole and their ancestral language (henceforth AL) for two or more generations. This allowed the creole to stabilize and expand before any children needed it as their own native language. If anything, rapid nativization would have imposed greater superstratal influence on the creole since native-born children generally have greater access to the local superstrate than foreign-born adults (Singler 1992:327-328).

These views stand in sharp contrast to the long-held hypothesis of Derek Bickerton that creoles form abruptly in a single generation. Bickerton (1981, 1984a) associates the pidgin-to-creole process with first language acquisition and claims that children acquire pidgins as first languages by activating an inborn language bioprogram: “The innovative aspects of creole grammar are inventions on the part of the first generation of children who have a pidgin as their linguistic input.” (1984a:173) Bickerton views pidgins as chaotic and unlearnable as native languages, and regards creoles as proof of the innate

nature of language. His work has been influential in general linguistics (e.g. Pinker 1994) and sometimes stands unchallenged by critics of the nativist position (e.g. Sokolov & Snow 1994).

Bickerton has often described the bioprogram as necessity-driven. It exists to provide humans with a native language if circumstances deny them their linguistic “birthright” (Bickerton 1976; 1984b; 1992). If children of captured slaves or immigrants manage to acquire their AL, the bioprogram would remain dormant unless their acquisition failed to instill native competence. So for creole formation to occur entirely in a single generation, Bickerton must argue that: (1) ALs immediately receded as targets for L1 acquisition, (2) ALs were not sufficiently acquired when they were available as L1 targets, (3) the target replacing the ALs was the pidgin spoken by parents. The following statement in Bickerton (1981:5) explicates these assumptions:

Since none of the available vernaculars [ancestral languages] would permit access to more than a tiny proportion of the community, and since the cultures and communities with which those vernaculars were associated were now receding rapidly into the past, the child born of pidgin-speaking parents would seldom have had any other option than to learn that rudimentary language, however inadequate for human purposes it might be.

It is taken for granted that ALs have little social value in the new land and that immigrants would rapidly abandon their native cultures and communities. Children born in the new social milieu are assumed to have acquired their pidgin from parents, and not some other source. Such acquisition is held to be essential to day-to-day life and not an option.

Bickerton based his unigenerational model of creole formation on HCE data published in Bickerton & Odo (1976) and Bickerton (1977a). He found that the language of locally-born speakers was more significantly elaborated than the speech of Japanese and Filipino immigrants who came to Hawai‘i in the early 20th century. These differences were mostly semantic and syntactic. His oldest locally-born informants were born prior to the arrival of these immigrants so he surmised that the pidgin they originally encountered was also unelaborated. Since the formation of HCE began no later than the turn of the century (Reinecke 1969), he regarded the elaborations as locally-born innovations.

These findings therefore stand as a key counterexample to the creoles described earlier. Since they constitute the strongest support for the bio-

program hypothesis, critics have questioned the validity of his evidence and the role of the locally born (Goodman 1985; Holm 1986). Since Bickerton's theories have had a strong impact on creolistics, acquisition research, and general linguistics, his scenario of HCE genesis deserves close inspection. But until now, no study has evaluated the Bickertonian model by assessing historical data on the course of nativization and creole formation in Hawai'i. That is the purpose of this paper.

3. Demographic Evidence of Nativization

The role of nativization in creole genesis is most commonly assessed with demographic evidence. Natality statistics, immigration reports, and general census data often provide information on the relative size of foreign-born and locally-born populations, ethnic and generational differences within populations, and the time when populations took shape and underwent demographic transitions. These data have linguistic implications. For instance Lefebvre & Lumsden (1989) and Singler (1992) argue that native-born populations contributed little to the formation of Haitian and the Surinamese creoles respectively, as statistics reveal them to have been minuscule in size throughout the early developmental period.

The "event hypothesis" of Baker (1982) and Baker & Corne (1986) provides an explicit framework for relating changes in the population to linguistic development. It highlights important demographic transitions (Events) that redefine the speech community. The founding of the colony may be designated Event 0. Event 1 occurs when the proportion of substrate speakers (slaves or immigrants) equals the proportion of superstrate speakers (most commonly masters or members of the upper class). The linguistic consequence is reduced access to native speakers of the lexifier, causing increasing numbers of new arrivals to learn the language from other substrate speakers. This enables the emerging contact language to stabilize as a system distinct from the lexifier. The longer the interval between Events 0 and 1, the more the contact language would resemble the superstrate. As for locally-born children, they would acquire the superstrate as well since relations with the elite are generally more intimate among the locally born than the foreign born (Singler 1992). The locally born would thus control a continuum ranging from the contact language to the superstrate.

Event 2 occurs when the native locally-born population equals the population of superstrate speakers. This reduces their access to the superstrate and allows their version of the pidgin to jell (stabilize) as a separate language. This would have already happened to the pidgin of the foreign born in the interval between Events 1 and 2. Until Event 2 the locally born would not have a significant effect on the language of the foreign born, aside from amplifying superstratal influence.

Event 3 occurs when the importation of foreign-born slaves or laborers ends, slowing the rate of growth of the locally-born population. This enables the population of superstrate speakers to catch up with the locally born. If this happens before the language of the locally born has stabilized (stabilization being delayed until after Event 2), it would probably never stabilize and the continuum would persist. The longer the interval between Events 2 and 3, the more likely the resultant creole would develop further as a separate language.

This framework was developed to explain why Réunionnais, a variety of French spoken in the Indian Ocean colony of Réunion, resembles non-creole French to a much greater extent than the nearby creole of Mauritius. As Baker & Corne (1986) show, the interval between Events 0 and 1 was about 50 years for Réunionnais but less than a decade for Mauritian. Thus Réunionnais had a firm superstratal base in the beginning. Then nearly 90 years transpired between Events 1 and 2 in Réunion, as opposed to 53 years in Mauritius. This delayed the stabilization of Réunionnais much longer. Finally only 5 years intervened between Events 2 and 3 for Réunionnais, while 30 years separated the two Events for Mauritian. As a result Réunionnais lacked enough time to separate itself from French. While the locally born were influential in Mauritius, they grew too slowly in Réunion.

Despite its usefulness in probing the development of these creoles, the event hypothesis runs into difficulties when applied to the Hawaiian situation. It was designed for colonies which lacked indigenous populations. The presence of a substantial indigenous population in Hawai'i creates two problems for an event analysis. As non-white substrate speakers, native Hawaiians vastly outnumbered Anglophone whites in the early 1790s, the time of Event 0. But by virtue of being indigenous Hawaiians they were also locally born. This fact blurs the difference between Events 1 and 2 and makes all three Events simultaneous. Event 3 is of no relevance since labor immigration did not commence for many decades after these Events.

One possible solution would move Event 0 up to 1876, the real start of the

sugar industry. This is when major labor immigration began. But Hawaiians still outnumbered both immigrants and whites. The 1878 census listed 44,088 Hawaiians and 3,420 Part-Hawaiians, as opposed to 3,262 whites. The numbers of non-Hawaiians increased until they surpassed Hawaiians in 1884, when the census listed 40,014 Hawaiians out of a total population of 80,578. Whites did not catch up with them until the 1930 census (38,268 Hawaiians versus 44,895 whites).

Another solution would exclude Hawaiians from the analysis. This would be undesirable since Hawaiians surely played a key role in the formative processes of HCE, but at least it would shed light on the transitional patterns of the immigrant foreign-born and locally-born populations. The results are displayed below in Table 1, with data from Réunionnais and Mauritian for comparison:

Table 1. Demographic transitions in Hawai‘i, Réunion, and Mauritius

	Hawai‘i		Réunion		Mauritius	
	Date	Years	Date	Years	Date	Years
Event 0	ca. 1790	—	1663	—	1721	—
Interval	—	82	—	52	—	6
Event 1	b. 1872	—	1715	—	ca. 1727	—
Interval	—	24	—	90	—	53
Event 2	b. 1896	—	1805	—	1780	—
Interval	—	39	—	5	—	30
Event 3	ca. 1935	—	1810	—	1810	—

Event 1 occurred shortly before 1872 because the census for that year mentions 2,846 non-whites (excluding Hawaiians and Part-Hawaiians) and 2,520 whites (Nordyke 1989). So even before the massive immigration of 1876 had commenced, there were more non-whites than whites.³ The interval between Events 0 and 1 lasted much longer than the same interval in Réunion, and the social situation in Hawai‘i was similar to the *société d’habitation* phase which Chaudenson (1992) posits for early Réunion: small-scale agriculture, relatively close contacts between whites and non-whites, and low proportions of non-whites. The demographic situation of this period differed notably from the one documented for Mauritius.

Figures for the locally born may be obtained by subtracting the number of foreign born from the overall non-white population.⁴ In 1896 the non-white and non-Hawaiian locally-born population numbered 13,733, as opposed to

7,247 whites. The interval between Events 1 and 2 was at most 24 years, half and one-fourth as long as the intervals in Mauritius and Réunion respectively. Stabilization was therefore delayed the least for HCE, the most for Réunionnais, with Mauritian falling somewhere in between.

Hawai'i also diverges markedly from Réunion in the timing of Event 3. Immigration to Hawai'i never really came to an end, but the importation of plantation labor fell sharply after the mid-1930s (Reinecke 1969; Nordyke 1989). This makes the interval between Events 2 and 3 at least 39 years, one-fourth and eight times longer than the intervals in Mauritius and Réunion respectively. So HCE had the most time to stabilize and expand, Mauritian had slightly less time, and Réunionnais had the least amount of time. This predicts that HCE developed as the most radical (distinct from the superstrate) creole of the three, all else being equal.

Bringing our attention back to the native Hawaiians, what effect did they have on the stabilization process? I will show in section 5 that they shifted early from their AL to the developing contact language, ahead of certain other groups. Thus the size of the locally-born population undergoing language shift was likely much higher than indicated above for the early phases. Increasing the population size would shift back the date of Event 2 further and narrow the interval between Events 1 and 2. The event hypothesis, therefore, suggests that conditions were right for HCE to stabilize and develop as a distinct language among the locally born.

Singler (1986, 1992) gauges the rate of demographic nativization by comparing the proportion of the locally-born and foreign-born populations. The date when the two reach numerical parity would thus mark another important demographic transition. I will call this Event X. Singler (1992:326) indicates its significance in this way: "The faster the locally-born population emerges and the larger it is, the greater the influence of the lexifier language". This relation is problematic however because it does not take into account the relative size of the group speaking the superstrate lexifier language. If Event X occurs sometime after Event 2, the locally born would outnumber the foreign born as well as superstrate speakers. Table 2 displays the proportions of the locally-born and foreign-born groups at various census dates, again excluding Hawaiians and whites.

Table 2. *Non-Hawaiian locally-born and foreign-born populations*

	Locally born		Foreign born	
	N	%	N	%
1878	947	9.0	9,530	91.0
1884	2,040	5.6	34,306	94.4
1890	7,495	15.2	41,873	84.8
1896	13,733	19.8	55,783	80.2
1900	21,275	18.3	95,070	81.7
1910	47,936	31.3	105,426	68.7
1920	94,595	44.2	119,563	55.8
1930	163,657	51.6	153,819	48.4

These data show that Event X took place between 1920 and 1930, roughly 30 years after Event 2. At this point the locally-born group outnumbered both the foreign born and whites (in 1930 locally-born and foreign-born whites totaled 44,895). Since Event 2 presumably weakens the locally born's access to the superstrate, the significance of Event X to superstratal influence is uncertain. Its importance instead lies in determining the amount of involvement the locally born have with the foreign born. After Event X contact between the two groups would decrease, distancing the locally born from their ancestral languages.

The pace at which the locally-born group reached numerical parity with the foreign-born compares very favorably with the creoles discussed by Singler. In Jamaica only 10 percent of the slave population was locally-born by 1752, nearly 60 years after the locally-born population took shape. If we posit 1876 as the start date for Hawai'i's locally-born group (prior to then most locally born were whites, Hawaiians, or Part-Hawaiians), it took only 15 years for them to reach the same point. After almost a century of slavery, the percentage reached 21.6% in Jamaica. The same development took 34 years in Hawai'i. Thus the rate of nativization, as defined by Singler, was rather rapid for HCE.

However Table 2 also shows that the rate was not smooth throughout the 50-year period. There was a slight dip in the beginning to 5.6%, and then it remained fixed in the 15-20% range in the 1890, 1896, and 1900 enumerations. But in 1910 the rate suddenly rose to 31.3% as the locally-born population doubled, and it climbed steadily thereafter. Therefore the first decade of the 20th century marked a transition between periods of stasis and change in the ratio between the two groups. I believe this was a crucial transition as further evidence will show.

While it is interesting to study demographic transitions globally, closer analysis of variation within the locally-born population will produce a more nuanced account of the nativization process. The most salient factors were ethnicity (or language group) and generation. The major ethnic groups brought to Hawai'i as labor were Chinese, Portuguese, and Japanese. The Chinese were the first group to respond to the call for labor; in the 1872 census they numbered only 1,938 but by 1878 they dominated the foreign-born population. Portuguese immigration began in 1878 and by 1884 they accounted for over a quarter of the foreign-born population.⁵ In 1885 Japanese labor was introduced, and by 1896 they surpassed the Chinese as the most numerous foreign-born ethnicity. More than half of the immigrants residing in Hawai'i were Japanese between 1900 and the 1920s. Political changes in 1900 forced an end to Chinese immigration; their numbers thereafter dwindled. Portuguese immigration ended in 1888, only to resume for a brief period between 1906 and 1913.

Later groups included Koreans (who arrived in 1904 and 1905 during the Russo-Japanese War) and Filipinos, the last major group lured to Hawai'i's plantations. They were introduced in 1907 and numerically surpassed foreign born Japanese by the 1930 census. They arrived too late to figure significantly in the nativization process, as sociolinguistic evidence will demonstrate.

Generational differences are not readily apparent in demographic data and must be deduced with additional evidence. I will adopt the system proposed by Corne (1994, 1995) for labeling generations: G1 refers to the non-native foreign-born generation, G2 refers to the first locally-born generation (children of immigrants), and G3 refers to the second locally-born generation (grandchildren of immigrants). Generational evidence is highly important in view of Bickerton's claim that G2 children alone invent creole grammar.

Table 3. Ethnic composition of the locally-born population

	Chinese		Portuguese		Japanese		Filipino	
	N	%	N	%	N	%	N	%
1878	129	2.1	50	10.3	—	—	—	—
1884	315	1.7	590	5.9	—	—	—	—
1890	1,451	8.7	4,117	32.4	250	2.0	—	—
1896	2,234	10.3	6,959	45.8	2,078	8.5	—	—
1900	4,026	15.6	10,604	58.0	4,877	8.0	—	—
1910	7,188	33.2	14,716	66.0	19,875	24.9	—	—
1920	12,343	52.5	21,208	78.5	48,584	44.5	2,303	11.0
1930	19,702	72.5	23,875	86.5	91,206	65.3	10,380	19.7

Table 3 splits the locally-born figures into major ethnic categories for the relevant period (percentages are of the total population of the ethnic group present at the time).

These figures reveal significant variation among ethnic groups. Although the average proportion of locally born was 19.8% in 1896, the figure was twice as high for the Portuguese. Event X had occurred by 1900 in the Portuguese population, almost 30 years before parity was reached overall. This suggests that the Portuguese were the fastest group to nativize. Event X next occurred in the Chinese locally-born group shortly before 1920. Japanese rank third with Event X occurring sometime in the 1920s. By 1930 all groups outnumbered their foreign-born counterparts except for the Filipinos, the last of the major ethnic groups to arrive. If these figures are indicative, they suggest that the Portuguese were the first to switch from their AL to the emerging creole, followed in succession by the Chinese, Japanese, and lastly, the Filipinos. Sociolinguistic evidence in sections 4 and 5 will reveal exactly the same pattern.

As far as generational factors are concerned, census reports do not distinguish G2 from G3. However it is possible to make some inferences from known facts. The first members of G2 were born approximately a year after the start date of G1 immigration. Then G3 began roughly 17 years later, supposing 17-18 years as the average length of a generation (Corne 1995). On this basis the start dates for G2 and G3 are calculated for each ethnic group in Table 4:

Table 4. Proposed start dates for G1, G2, and G3

	Chinese	Portuguese	Japanese	Filipino
G1	1877	1878	1885	1907
G2	1878	1879	1886	1908
G3	1896	1897	1904	1925

Most G2 and G3 were born some time after these dates. The Portuguese G2 began in 1879 but five years later only 590 Portuguese were listed on the census as locally born.⁶ Similarly 1886 marked the beginning of G2 for the Japanese but again only 250 were born in the four years that followed. In contrast, 6,959 locally-born Portuguese were counted the year before the start of G3 and 4,877 Japanese had been born by 1900. So several years elapsed before the G2 population emerged on a large scale.

Table 3 shows that the Chinese and Portuguese G2 populations were minuscule in 1884 and that both experienced substantial growth in the following six years. Therefore these G2 populations arose in substantive numbers *ca.* 1887. In 1890 the Japanese G2 group was still minuscule. But by 1896 a definite population had taken shape. So the G2 Japanese probably became visible *ca.* 1893.⁷ While a few G2 existed as early as the dates in Table 4, most Chinese and Portuguese were born after *ca.* 1887 and most Japanese after *ca.* 1893. This means that the G3 Chinese and Portuguese population emerged substantially *ca.* 1904 and the G3 Japanese *ca.* 1910. These dates are rough estimates of course, but they accord well with generational data from autobiographies (see section 5).

Although the G2 Japanese population expanded in the 1890s, Table 3 shows that it was dwarfed by tremendous increases in the next decade. In 1910 the census listed 19,875 Japanese as locally born, four times the number reported in 1900. So three out of every four non-immigrant Japanese was under the age of 10 in 1910. This development shifted the balance between the locally-born groups. In the 1890s G2 Portuguese greatly outnumbered Chinese and Japanese locally born (6,959 Portuguese versus 2,078 Japanese in 1896) but by 1910, they were surpassed by the Japanese.

The Japanese baby boom of the early 1900s was caused by the massive immigration of 34,000 Japanese in the four years preceding 1900. The 1900 census reported 56,234 foreign-born Japanese but only 12,360 of these were listed in the 1896 census (Schmitt 1977, Nordyke 1989). So the swift movement of G1 labor into Hawai'i sent ripples of rapid growth through the G2 and G3 generations. Natality statistics show how fast the birth rate accelerated: 221 Japanese babies were born in 1900, 486 in 1901, 590 in 1902, and then for the next four years the rate remained in the 600-700 range (Nordyke 1989: 209). But in 1907 it doubled to 1,487, then to 2,347 in 1909. Thus ten times more Japanese babies were born in 1910 than in 1900.

If the G2 baby boom began in earnest by 1907, as indicated by natality statistics, a similar increase in the G3 population probably occurred *ca.* 1924.

Demographic data highlight the 1900-1910 decade as pivotal in the growth of the locally-born population. Prior to 1900 G2 Chinese and Portuguese dominated. Very little if any G3 had been born yet. In 1910 the largest proportion of the population was G2 Japanese; G3 Chinese and Portuguese had also emerged. So the complexion of the population changed considerably during the decade.

Before concluding this section, it would be helpful to review some key points which may have been obscured by complexities in the data. If the assumptions of the event hypothesis are valid, HCE probably began to stabilize in the 1890s. It had at least 40 years to stabilize and develop as a language distinct from the lexifier. The locally-born population grew fast enough to surpass native English speakers by 1896 and the foreign born by 1930. G2 began substantially around 1887 for Chinese and Portuguese locally born and 1893 for the Japanese. G3 had emerged by 1904 in the Chinese and Portuguese population and 1910 in the Japanese population. There was an explosive increase of G1 Japanese between 1896 and 1900 which accelerated the growth of G2 in 1907 and (probably) G3 in 1924. The 1900-1910 decade was transitional between a locally-born population lacking G3 and consisting mostly of G2 Portuguese and Chinese and a later population which incorporated G3 Portuguese and Chinese and was dominated numerically by G2 Japanese.

4. Sociolinguistic Evidence of Nativization

Most studies rely heavily on demography because of the scarcity or nonexistence of data on sociolinguistic factors of creole formation and nativization, such as the status of ALs and the pidgin/creole in the home and community, relations among ethnic and social groups, and attitudes towards the pidgin/creole, ALs, and the superstrate. The existence of such evidence for HCE offers a rare glimpse into the social context of language learning and use during creole formation.

The importance of this evidence lies not only in its quantity but also in its diversity. In no other case of creole formation has anyone found such data from all three groups involved in the formation process: the superstrate elite, foreign-born speakers of substrate languages, and the locally born. In most cases only members of the first group provide comments on language learning and use.

In the present corpus, educators supply most of the data for the superstrate group. They sometimes denigrate the pidgin/creole or at least cast it in an unfavorable light. Comments in teacher memoirs, newspaper and journal articles, and reports of the Board of Education may contain negative evaluations, though most writers made informative or even astute observations. As

an example, consider the following remark by Weimer (1934:57):

Recently, one of the principals of another intermediate school conducted a straw vote among 73 of his pupils to find out why they did not use better English. Thirty-three mentioned fear of ridicule, 17 alien home language, 13 claimed a lack of practice, while 10 said they did not try.

This provides useful information on factors favoring the use of HCE by children. But the same individual wrote a few pages earlier: “ ‘Pidgin’ English is fragmentary, unintelligible, and in a larger sense, irrational. . . Irrespective of race, tongue, or creed, it spreads like some contagious infection” (50-51). So one must use discretion in interpreting such comments. Fortunately this is an extreme case and not representative of most statements by educators. Also the formation and development of HCE occurred under the observation of linguists and professional sociologists, ensuring data tainted less by racism and social stigma than remarks by laypeople or propagandists.

My database also contains remarks by foreign-born Japanese, the largest of the substrate groups. These were published in Japanese-language newspapers (*Yamato Shimbun*, *Nippu Jiji*, *Hawaii Hochi*, and others) and memoirs of community leaders. Of considerable interest are the writings of Takie Okumura, who came from Japan in the 1890s and founded the first Japanese language school in Hawai‘i. These sources for the most part do not distinguish the pidgin/creole from Standard English and are primarily concerned with the survival of the AL in the locally-born Japanese population.

Finally, and most significantly, sociolinguistic comments by locally-born children exist and furnish extensive information on language use and attitudes. This evidence will be considered in section 5.

In the discussion which follows and in the rest of the paper, I will use “PE” as a general label for pidgin and creole varieties of English. This avoids making a slippery distinction between the two when written sources use the same word (pidgin) for both.

Perhaps the most striking pattern in the data is a disparity between language use in the late 19th and early 20th century (particularly after 1910). The first decade of the 20th century was apparently transitional between an early period of widespread multilingualism and AL maintenance and a later period when PE has supplanted ALs in child-parent and peer group discourse.

Early sources show that G2 children usually acquired their ALs in the 1880s and 1890s. The 1886 *Biennial Report of the Board of Education* states: “The child learns his vernacular by the easiest and sweetest of methods —

from a mother and father” (Anonymous 1886:4-5). James H. Blount, a member of the U.S. Congress sent to Hawai‘i to investigate political and social conditions, noted that in Honolulu “the groups of children playing along the streets use their native tongue” (Blount 1893:19). The *Report of the Minister of Public Instruction* in 1897 declared: “To a large extent children use the language of their parents while at home” (Anonymous 1897:7).

But some writers suggested that children lacked full competence in their ALs. An article in *The Friend* (December 1891, p. 99) commented that “[t]he Portuguese born here are . . . speaking their mother tongue imperfectly” and the *Thirty-Third Annual Report of the Hawaiian Evangelical Association* claimed that Japanese children “have been growing up in this polyglot city [Honolulu] without a correct knowledge of their own tongue” (Anonymous 1896:34-35). Of course it is not clear what the deviations were (in the case of the Japanese children, perhaps the writer only meant an unfamiliarity with polite forms or the Japanese writing system), but these remarks could indicate that the process of language shift had already begun.

While children in urban areas formed peer groups outside of school (as Blount observed), the situation was often different on plantations. Since laborers and their families were segregated ethnically (Reinecke 1969), children did not socialize to a great extent outside of their ethnic group until they entered the public schools. Communication problems were typical at the start of the school year:

The Hawaiian with his twelve-lettered alphabet, the Chinese boy who knows not the sound of *r*, the Japanese whose vernacular has no *l*, the German and Portuguese who are ignorant of the vocal or aspirate *th* — all these nationalities go to the same school on Monday, the opening day, not one able to communicate with the other, and the teacher speaking an entirely different language from any of the pupils (Anonymous 1886:7).

Teachers sometimes coped by learning their students’ ALs. Miss Turner, a teacher on the Island of Hawai‘i, “has mastered Hawaiian and is rapidly mastering Portuguese, and therefore is able to fix in the minds of her pupils the English equivalents of the words familiar to them in their parent tongue” (*Saturday Press*, 4 April 1885, p. 3).

ALs formed the basis of social interaction, children mingling by language group. The *Pacific Commercial Advertiser* (16 January 1888, p. 2) described the situation in a school in Lihu‘e, Kaua‘i:

At recess the teacher called my attention to the variety of languages used on the playground, saying that if the curse of Babel had fallen on the school their speech could not have been more confounded. I looked, and there was a small group speaking German, some speaking Norwegian, a company talking Portuguese, and a great number using the Kanaka language; while English was stopped like the study of Greek when the lessons were over.

Wider social relations could only be established by learning other ALs. Multilingualism was therefore the norm; children were commonly trilingual (Prescott 1899:13) and “readily learn two, three or even four languages, most of them conversing well in three. The Portuguese and Chinese have three languages at their tongues’ end always. The Hawaiians speak two, and also have command of some very sharp Chinese phrases when they get into altercations with the children of Far Cathay. . . [T]he child easily acquires speech foreign to its mother tongue” (*Paradise of the Pacific*, October 1901, p. 16).

Many writers noted that children (primarily G2 Japanese) code-switched between their AL and other languages used in the community. According to Anonymous (1896:34-35), “some Japanese children are heard in the streets to speak a mixture of English, Hawaiian, Portuguese, and Japanese”. A speech sample was provided:

- (1) *Mama nuinui hana, ko nai.*
 ‘My mother had a lot of work and could not come.’

This sentence contains English *mama*, Pidgin Hawaiian *nuinui* ‘a lot, very’ and *hana* ‘work,’ and the Japanese phrase *ko nai* ‘come NEG-PAST’.

The *Yamato Shimbun* (26 May 1898, p. 2; translated from Japanese) noted that “[t]he children use a strange language mixed with Chinese and Japanese”. Takie Okumura, when he first arrived in 1894, met a locally-born Japanese girl but “could not understand what she was saying. When I asked Mr. Ishimura, he explained to me that the little girl was using English, Kanaka [Hawaiian], and Japanese, and that is how the children speak in Hawaii” (*Nippu Jiji*, 1 January 1915, p. 8; translated from Japanese).

So locally-born children prior to 1900 were multilingual and did not depend on PE in their linguistic interaction, though it undoubtedly formed part of their linguistic repertoire. Prescott (1891:98) observed its use in the classroom: “If a boy was wanting a pencil ‘me lend!’ ‘me, too lend!’ could quickly be heard”. Another teacher noted in 1896 that Portuguese children “spoke pidgin English with that inimitable Portuguese accent” (McKay 1896-1901: 100). In court testimony, John Vierra represented his son as speaking PE

(4CR-869, 9 February 1897, p. 1). None of these sources, however, indicated that PE was any more important than ALs in parental or peer interaction.

The situation changed considerably in the following decade. While earlier writers said that parents and children generally spoke their ALs at home, others now claimed that children were using some form of English in this context. The *Yamato Shimbun* (31 October 1905, p. 3; translated from Japanese) stated:

Generally speaking, Japanese children seem to speak 70% English and 30% Japanese. When their parents speak Japanese to them, they are likely to answer in English except when they think the English is too hard for their parents to understand. Of course, the children use only English among themselves. They still use some Japanese terms to express anger or excitement, but those words also are getting replaced by English ones.

This quote significantly asserts that children, not adults, introduced English into parent-child discourse. This is contrary to Bickerton's assumption that parents first used PE, which their children later acquired. It is also important because it claims that PE dominated in discourse between siblings. This marks a clear shift away from ALs and multilingualism.

Outside the home, peer relations began to involve less multilingualism and more dependence on a single language, PE. A decade after Blount observed the general use of ALs by children in Honolulu, Wallace R. Farrington (later elected Governor of Hawai'i) wrote: "However pure the English, Chinese, Portuguese, or Hawaiian they may speak in the school or homes, they have a complex pidgin English that is a universal language. They all meet on the 'I-bin-go' method of communication" (*POTP*, December 1904, pp. 43-44).

By the 1910s, PE was everywhere. Many educators lamented: "What can be done to lessen the use of pidgin-English among the schoolchildren of Hawaii?" (*Kohala Midget*, 31 July 1912, pp. 5-6). In striking contrast to the situation in 1888 (see above), PE was dubbed "the Esperanto of the school playground" (*Hawaii Educational Review*, February 1914, pp. 6-7). The general use of PE in peer groups was vividly portrayed in a short story by Breckons (1915:12):

[There] stood an admiring group of little boys of all colors and nationalities. "Look, Tony, look. One cow-boy picture. Look!" The little Hawaiian pointed excitedly to a multi-colored poster representing a band of impossible Indians fleeing from some equally impossible cow-boys. The boy addressed, a dirty

Portuguese lad of about twelve, turned his attention to the poster designated, and his eyes grew round as he exclaimed, "Oh look, look! The Indian he got one gun. He been shoot one cow-boy."

Some writers even claimed that the social use of PE promoted linguistic change, most notably a reduction of variation in the use of particular features (errors) across locally-born ethnic groups. The *Pacific Commercial Advertiser* (18 September 1919, p. 2) stated: "The children have imitated each other in their close association so that no error is confined to one nationality". Adams (1925:1), a sociologist at the University of Hawai'i, keenly observed:

Where the national groups live in close community, the children play together to some extent and the common language is English brought from school by the older children and taught to the younger naturally in play. . .As the children attempt to use the language away from school they fall into pretty serious errors and tend to standardize these and pass them on to the younger children. There is not only much miscommunication but an erroneous construction with characteristics of plantation pidgin.

This quote is important because Adams identifies older children, not parents, as the source of the PE acquired by young children. Bickerton, as noted previously, assumes the source to have been parental. Adams also claims that older children, in their social interaction, "standardized" particular features, which were later acquired by younger speakers.⁸ This statement seems to describe the stabilization process. Adams' suggestion that children derived structural features from both plantation pidgin and the superstrate is also noteworthy, in view of the prediction by the event hypothesis that superstratal influence would be slight.

Although school-educated children were bilingual in PE and the superstrate, they rejected Standard English in peer group relations. This was partly due to ethnic and socioeconomic tensions, as "the language of the schools" was essentially the language of the white bourgeois minority. It was esteemed as a means of securing middle-class employment by young adults, but younger children found no practical use for it outside school. Adam Ornelles, a student from McKinley High School, remarked: "As soon as they step out of the school, they speak a shameful pidgin English, and if they are corrected by some other student, they answer that they are not in school now, and they should not worry" (PCA, 13 August 1915, p. 2). Another writer astutely noted:

It is a common experience that pupils even of the higher grades will revert to the lingo of the street as soon as they get out of the environment of school.

This is natural, for it is a second mother tongue to them and one in the use of which they can throw off the restraints of school, just as they don their working clothes as soon as they reach home. . . [I]t is the fashion to speak this short-cut jargon, and any attempt to speak the language of the educated people on the street, in the home or on the playground, will bring down upon our pupils the ridicule of their playfellows, and often of parents as well (*HER*, September 1921, pp. 2-3).

Their linguistic attitudes and “sensitive pride” in speaking PE (*PCA*, 9 April 1915, p. 2) suggest that the locally born valued PE as an indicator of group identity. This adds further evidence that the linguistic situation had changed after 1910.

Although PE was indispensable to school-aged children in their social relations, many continued to acquire their ALs from parents, particularly in plantation communities: “[T]he mother tongue of its particular nationality is practically the only one used in a given camp. In these homes parents and children converse in their own language almost exclusively. . . Hence they often enter school at six or seven years of age knowing scarcely a word of English” (*HER*, September 1921, p. 1). This situation persisted through the 1920s at least. In contrast, Bickerton believes that ALs would have “receded rapidly into the past” on plantations, causing locally-born children to learn PE from parents.

But while some families were conservative in retaining ALs, an increasing number used PE exclusively. Adams (1925:1) noted that PE was “the family language of some Chinese families — the only language known to both parents and children”. According to Brigrance (1938:20), “Pigeon English is essential, even a necessity, so that children may speak to their parents”. As a result, “it is not unusual for a child and his grandparents to be mutually not understandable except in commonplaces” (Jones 1942:158). Interestingly, such statements are typical of the 1920s and 1930s, not of any earlier period. Bickerton, on the other hand, assumes that G2 children were in this situation all along.

Kaapu (1937:25-26) described the differing use of PE in the home among the various ethnic groups:

In the case of the Portuguese, the home language seems to have become entirely English, and it is chiefly so for the Hawaiians. Even where the Hawaiian parents use Hawaiian, the younger generation can usually understand it only superficially, and speak it scarcely at all. All Filipinos reported the home language as both English and Filipino, English usually the first. . . The Chinese children make very little use of Chinese, and understand it only to a limited extent.

Japanese children however were bilingual and occasionally monolingual in their AL. The Portuguese were monolingual PE-speakers, the Chinese and Hawaiians made only a very limited use of ALs, Filipinos spoke their AL but less than PE, and the Japanese used their AL most of all. This is almost the exact same pattern revealed by autobiographical data (see section 5). Most significant is the fact that Portuguese and Chinese children, due to the cessation of G1 immigration, were almost entirely G3 by the 1930s. Japanese children, however, were largely G2 due to continued G1 immigration in the 1920s and 1930s. This suggests that the exclusive use of PE between parents and children was largely a G3 phenomenon, *contra* Bickerton.

Bickerton's scenario of nativization, sketched at the outset of this paper, bears little resemblance to what actually happened in Hawai'i as documented by contemporary observers. ALs did not lack social value when G2 children were born in the 1880s and 1890s; they were essential both in and outside the home. Because of multilingualism, ALs permitted access to more than "a tiny proportion of the community". ALs and immigrant cultures did not rapidly vanish; plantations fostered ethnic communities and parental use of ALs, at least through the 1920s. Parents did not initiate the use of PE in the home, as a target for L1 acquisition; the actual situation was the reverse: school-aged children brought PE into the home, used it with younger siblings, and then with parents (in place of the AL, which they had already acquired). While many parents and children used PE exclusively in the 1920s and 1930s, this was a late development and one which involved G3 instead of G2 children.

The 1900-1910 decade stands out in particular as intermediate in these developments. Multilingualism among children, widely reported before 1900, was no longer mentioned by contemporary writers after 1910. The general use of PE outside the home and among siblings, the stigmatization of the superstrate, and the eventual use of PE by parents were all established features of the later period; very little if any evidence of these can be adduced before 1900. Children acquired ALs during both periods, but in the second, only G3 children were inclined to learn PE exclusively (G3 children largely did not exist before 1900, as suggested above). So the same period which was transitional demographically was also transitional sociolinguistically.

John Reinecke, writing in the 1930s, came to the same conclusion. He indicated the same decade as transitional on the basis of age differences among HCE speakers in his day:

For some time previous to 1920 there must have been a considerable number who entered school able to converse in English. Of those speaking dialectal English [HCE], a good many are between the ages of thirty and forty. Therefore, it was the present writer's deduction that the use of the English language at home and on the playgrounds became fairly general in the second decade of the present century, and that by 1925 the children were (with allowance due their age) as a general thing speaking the Island dialect [HCE] than very broken creole [foreign-born pidgin]. This estimate allows about one generation (1876 to 1900) for the formation of the creole dialect and its diffusion to the point that the average child would have more or less opportunity of hearing it from neighbors and possibly from elder children who had passed through the English schools, and another generation (1900 to 1925) for the schools to have bred a generation of young people to whom English was no longer a foreign language and who would use dialectal English in conversation with the younger children. The use of Island dialect must have been under way by 1900 and well established by 1910. (Reinecke 1969:166)

Sociolinguistic remarks in the autobiographies of children born during this transitional period comprise our next line of evidence.

5. Sociolinguistic Evidence from Autobiographies

In 1926 sociologist William C. Smith distributed a questionnaire to most junior high and high schools in Hawai'i. It contained over a dozen questions relating to when and where the students were born, how large their families were, who was recognized as the head of the family, the occupations and educational background of the students' parents, the ethnicity and places of origin of their parents, their attitudes towards their ancestral culture and customs, their relations with people of different ethnic backgrounds, their political identity as Americans, and their future plans. The questionnaire also included two questions of linguistic interest: "What languages do your parents speak and which are used in the home?", and "How do you value your native language?" Teachers gave copies of the questionnaire to students as homework with the assignment of composing their own life histories. Smith also taught a course on sociology at the University of Hawai'i and assigned a similar exercise to his students. This allowed Smith to sample from a large age range, between the ages of 13 and 24. His informants therefore were born between 1902 and 1913. Smith returned the following year and collected additional data.

The two linguistic questions could not have better suited the requirements of this study. They ask for precisely the information needed to test models of nativization and creole genesis: Did parents use PE with their children instead of ALs? How prevalent were ALs in the home? Were there differences in the way children would speak to their parents than siblings? What attitudes did children have towards their ALs? Did they regard PE instead of the AL as their native language? Did they speak their ALs well enough to consider them useful? All these questions receive answers in the life histories.

Since all the autobiographies were composed from the same questionnaire in response to the same questions, they lend themselves to quantitative analysis. Sample size would not be a considerable problem; the autobiographies occupy 17 rolls of microfilm, containing an estimated 1,550 life histories. Not everyone replied to all the questions, but all were answered by a large majority. The most impressive aspect of these data are the non-linguistic questions. These should enable us to correlate social and geographical factors with linguistic phenomena. It is also possible to study the autobiographies at a microlevel of analysis, from household to household, since they were composed by individuals.

The autobiographies can also be considered more reliable than oral histories recorded late in life. As teenagers and young adults, the situations and events from childhood were still fresh in the informants' minds. Smith was able to capture these memories only a decade later and in many cases the linguistic situation still persisted unchanged.

No data source is ideal, and the autobiographies are not without problems. The most obvious is the context in which they were written. Students did not compose them as personal memoirs for their own use. Instead they were solicited by teachers for a prestigious man most of them had never met. So the pressure to project a pro-American and pro-English language attitude was probably felt by the writers. While they were freely explicit in mentioning their parents' acquisition of PE, only rarely did they admit to their own use. In almost every case they said they spoke "English", without specifying further. This lack of explicitness does not devalue the data (since in either case the reference would have been to PE, aside from the few who spoke Standard English at home), but the reader should be aware of this tendency.

There may also be sampling biases in quantified analyses. While the corpus itself is quite large, it was drawn only from students who attended public schools. While this accounted for the vast majority of locally-born

youth, some were unable to attend school (Lydgate 1919). Even for children attending schools, the questionnaire may have systematically missed children who skipped the assignment because of writing insecurities or other reasons. So the Smith corpus does not represent the locally-born population evenly, though these biases were probably of minor significance.

Today the University of Oregon Library retains the original autobiographies and microfilmed copies are kept at Hamilton Library, University of Hawai'i. As preliminary fieldwork, I examined a total of 240 autobiographies. They represent students from the University of Hawai'i, the Territorial Normal School, Maui High School, McKinley High School, Hilo High School, Honokaa Junior High, and Laupahoehoe Junior High. The sample thus includes students from the three major islands of O'ahu, Maui, and Hawai'i and covers the entire age range. The sample includes substantial numbers of Japanese (N=128) but smaller quantities of Hawaiians (N=24), Chinese (N=40), Portuguese (N=33), and other groups (N=15). So some of the patterns suggested by autobiographical data may not be statistically significant. However since the entire corpus dwarfs the present sample, further research will produce much firmer results.

The questionnaire asked students to state their parents' ethnicity and place of birth, and the answers to these questions determine each student's ethnicity (relevant in cases of mixed marriage) and generational status. Together these data reveal overall patterns in the proportions of G1, G2, and G3 children across ethnic groups.

Table 5. Generational variation among Chinese, Portuguese, and Japanese

	Chinese		Portuguese		Japanese	
	N	%	N	%	N	%
G1	1	3.0	1	3.7	2	2.0
G2	25	75.8	11	40.8	95	95.0
G2/3	7	21.2	9	33.3	3	3.0
G3	0	0.0	6	22.2	0	0.0

In Table 5 neither Chinese, Portuguese, nor Japanese had significant proportions of G1 children. As for G2 the Japanese were highest at 95.0%, the Chinese ranked second at 75.8%, and the Portuguese had the lowest proportion at 40.8%. G2/3 refers to children who had both G1 and G2 parents. In many cases the G1 parent had immigrated as a child. Both the Chinese and

A portion of these informants also gave the date of their birth. This information allows us to compare the distributions of G2, G2/3, and G3 over the 12-year period. The results are not statistically significant for the Chinese and Portuguese because of the small sample size (N=29), but they fit into the expected pattern. Figure 1 below presents the distributions (light shading for G2, medium shading for G2/3, and dark shading for G3):



Now we turn to the linguistic evidence. In answer to Smith's first question, "What languages do your parents speak and which are used in the home",

students could have mentioned their ALs, English, or the use of both. I have classified these responses into two categories: (1) Students who reported any use of their ALs (N=165), and (2) Students who reported any use of English (N=102). Students who mentioned both their ALs and English belong to both categories. Most of these indicated which of the two prevailed in parent-child discourse (N=98). Table 6 presents the results by ethnicity:

Table 6. Variation in the use of ALs and English in parent-child discourse

	Hawaiians		Chinese		Portuguese		Japanese	
	N	%	N	%	N	%	N	%
Use of AL	17	77.3	36	94.7	21	72.4	91	100.0
Use of English	20	100.0	26	81.3	27	96.4	29	50.9
AL-dominant	4	23.5	10	52.6	2	9.5	38	92.7
English-dominant	13	76.5	9	47.4	19	90.5	3	7.3

The data show that the Japanese scored the highest in the use of ALs (100.0%), followed by the Chinese (94.7%), Hawaiians (77.3%), and Portuguese (72.4%). This pattern conversely mirrors the order in which locally-born ethnic groups surpassed their foreign-born counterparts in size (see section 3), and indicates that Hawaiians were linguistically similar to the Portuguese. It is noteworthy that the use of ALs was above 70% for all four groups. This suggests that ALs were generally acquired to some degree between 1900 and 1910.

The same order occurs with respect to English use. Hawaiians (100.0%) and Portuguese (96.4%) ranked first, then Chinese (81.3%), and lastly Japanese (50.9%). This pattern shows that Hawaiians and Portuguese were fast to nativize PE, Chinese were slower, and Japanese were the slowest. This is exactly the order predicted by demographic evidence. Interestingly the range of variation was greater for English use. While most informants (72.4-100%) reported some use of ALs, only half of the Japanese used English.

Lastly, Portuguese (9.5%) and Hawaiians (23.5%) had the smallest proportion of informants claiming to use more AL than English with their parents, Chinese were intermediate (52.6%), and Japanese by far had the highest proportion (92.7%). Again we find the same pattern. Taken together, the data in Table 6 indicate that strong differences existed between groups in their nativization of PE. They also show that Hawaiians patterned like the Portuguese (something demographic evidence could not reveal).

Of course these are broad generalizations and the actual situation was much more complex. Turning to an individual consideration of the autobiographies, we find a good deal of variation within each group. In the case of Hawaiians, some reported infrequent use of English. James Palaika, born in 1910, wrote: "I use the Hawaiian language when I am speaking to my parents ... They don't speak English so often. We use both languages in our home" (HJ-46, pp. 2, 7). In other cases Hawaiians would report speaking English when their parents spoke Hawaiian, such as this informant born in 1909: "My parents talked the Hawaiian language most of the time but speaking English fluently when they used it. Although he [sic] talked Hawaiian to us we spoke English to him. Sometimes he used to get mad and tell us to talk Hawaiian to him" (HH-56, pp. 1-2). Other informants said their parents spoke more English than Hawaiian. A student born in 1912 remarked: "The most common language used in our home was pigeon English. The reason why it was used was because my parents were not well educated and were unable to speak the English language well" (H-20, p. 3, 6). Another Hawaiian wrote: "English was used in our home most of the time but sometimes my parents spoke Hawaiian, when it was necessary" (N-188, p. 1).

In contrast to Bickerton's assumptions, parents spoke PE alongside their ALs as the following excerpts indicate:

They [my parents] understood the English language fairly well but couldn't speak fluently. In the home, we used the 'pidgeon' English and sometimes, Hawaiian was used in conversation with the older folks (H-20 [Hawaiian born in 1902], p. 3).

My mother speaks to the children in "pidgin" English. The languages used in our home were both Chinese and English. . . I use the language of my parents very often in speaking with my parents (MK-61 [Chinese born in 1909], pp. 3-4, 9).

At home we talked the corrupted form of English and the tongue of our parents (H-44 [Chinese born in 1903], pp. 1, 3).

We do not speak Chinese at home unless we speak to our mother and that is part English, pidgin and Chinese (N-154 [Chinese born in 1907], pp. 1-2, 5).

My parents understand pidgin English mixed with Hawaiian. When they talked to us they mixed a little Japanese with English so that we would be able to understand them (MK-19 [Japanese born in 1910], pp. 1-3).

Yet the use of English among siblings was categorical, especially for Hawaiians, Chinese, and Portuguese, but also for many Japanese:

[My parents] spoke the Hawaiian language to us but in answering them we spoke English. When speaking to my sisters and brothers I usually spoke "pigeon English" (HJ-22 [Hawaiian born in 1911], pp. 2-3).

At home we children spoke English to each other, but it was not very good (L-18 [Hawaiian born in 1911], p. 3).

We usually speak English to my brothers, sisters and English speaking friends at home. My parents at times scolded us for always speaking English at home (MK-26 [Chinese born in 1909], p. 2).

At home we children usually talk English, but sometimes our parents talk Chinese to us (MK-20 [Chinese born in 1909], p. 2).

We children have to speak Portuguese when addressing him [their father], otherwise English is used in the house (N-112 [Portuguese born in 1904], p. 2).

By this time, all of us children were speaking English at home more or less. Naturally, my parents learned to cultivate an ear to what we were saying and began to understand although they could not speak themselves (H-49 [Japanese born in 1905], pp. 1-3).

When we grew older, we [children] became became more proficient in speaking the English language than the Japanese language. For us it is much easier to speak 'pidgin' English (H-39 [Japanese born ca. 1902-1904], p. 2).

I speak English when I am with my sister but sometimes I have to speak in Japanese because my parents do not understand what we are talking about at certain times (HH-53 [Japanese born ca. 1908], p. 2).

At home I use Japanese language to my parents, but to my sister and brother I use English (HH-55 [Japanese born ca. 1908], p. 2).

We speak Japanese to our parents but we use English when speaking between our brothers and sisters (MK-9 [Japanese born ca. 1910], p. 4).

There were constant arguments between my sisters and brothers. . . . They argued in English and since mother could not understand English well, I had to settle these troubles (MK-54 [Japanese born ca. 1912], p. 2).

So while the use of PE in parent-child discourse was highly variable, children generally spoke some form of English among themselves. This pattern suggests that children acquired PE from other (possibly older) children rather than from their parents, especially since their PE was often unintelligible to parents as H-49, HH-53, and MK-54 observed. MK-26 also shows that the use of PE by children in the home sometimes occurred despite their parents' disapproval. Bickerton's assumption that children learned PE directly

from parents or other adults does not receive support from these data.

While most children indicated that they had learned their AL to some extent, many expressed insecurity in their competence or described themselves as passive bilinguals:

I can't talk much Hawaiian but I can understand what they say (L-33 [Hawaiian born in 1909], p. 4).

I do not know enough Chinese to have experienced the value of the language (MK-46 [Chinese born ca. 1912], p. 3).

I understood only a few of my parents' jokes etc., because I can understand and speak only a childish Chinese language (L-6 [Chinese born in 1912], p. 3).

In our home we all understand the most simple things in our language but when it comes to speaking we stumble here and there until we finally make such a mess of it that we are laughed at (HJ-5 [Portuguese born in 1911], p. 6).

I never learned to speak the Portuguese language except a few baby words. It is impossible for me to hold a conversation all in Portuguese. I understand the language pretty well (LJ-21 [Portuguese born in 1911], p. 4).

Aside from the Japanese school room, a Japanese boy seldom speaks Japanese. Even at home many of the Japanese boys born in Hawaii speak a sort of Pidgin English to their parents (H-40 [Japanese born ca. 1902-1904], p. 1-2).

I cannot speak Japanese very well. At home I spoke Pigeon Japanese (as I call it), to my father and spoke English to my mother (M-19 [Japanese born in 1910], p. 1).

Many of these cases represent first language attrition and describe the situation in the late 1920s, not the early 1910s when children first learned their AL. They may have originally had native competence in their AL but intensive use of PE among siblings and other peers led to its eventual loss. However some, such as LJ-21, may have never gained competence in their AL.

The Portuguese appear to have progressed the most in the direction away from their AL and this is also reflected in their language attitudes. G2/3 and G3 informants sometimes expressed ardently negative evaluations of their AL:

The children should not learn the language because we are citizens and should forget our grandparents' native language and learn *our* native language (MK-31 [Portuguese born ca. 1912], p. 3).

People living in America should learn its language and not another foreign

language. My parents agree with me in this question (HH-71 [Portuguese born ca. 1908], p. 5).

Sometimes my mother talks to me in Portuguese language. I hate this. I have told her many a many times not to because this is an American country, not a Portuguese country (HJ-88 [Portuguese born ca. 1912-1913], p. 7).

These sentiments reflect political ideologies and suggest that PE was valued as an indicator of local identity. Only Portuguese informants expressed themselves this way; most others found some value in their AL.

Outside the home children depended on PE in their relations. Only a single student (born on Kaua'i in 1905) mentioned the need for learning other ALs: "During my early days I had for playmates children of different nationalities. Before I was eight years old I learned to speak the Porto Rican language and understood a little of the Japanese language" (N-114, p. 2). Most others acquired English from their friends in the community if no one else spoke PE in the home, a common situation in Japanese families:

The only language used at home was Japanese. It was after my seventh birthday that I was first taken to an English kindergarten by my sister. . . . I had already picked up simple English words and phrases through my association with children of school age (N-8 [Japanese born ca. 1905-1906], p. 2).

We children spoke good Japanese while we were small but as we grew older and conversed with the camp boys & girls we began to use bad Japanese, bad English, or pigeon English (LJ-64 [Japanese born ca. 1910], p. 4).

Standard English was stigmatized in peer relations and children who spoke inappropriately were sometimes ridiculed:

When I was in Central Grammar I know we were taught to speak correct English both in and outside of school but when I returned to Maui, my old friends made fun of my English which embarrassed me. They said I was "stuck up" and "you think you Haole [white]" so I had to use the pigeon English and I am having the hardest time breaking that habit (LJ-81 [Part-Hawaiian born in 1910], pp. 1-2).

As was the case with general sociolinguistic data, autobiographical evidence has contradicted Bickerton's scenario on a number of issues. Table 6 shows that ALs were spoken in a majority of households for all groups, suggesting at least that G2 children managed to acquire their AL. Of the immigrant groups, only the Portuguese used English more extensively than the AL in parent-child relations and they already included substantial numbers of G2/3 and G3 (Table 5). The Japanese, who had only a small number of G2/3

and no G3, used English the least in this context. Parents who did speak PE with children used their ALs as well. The highly variable use of PE by parents contrasts sharply with its widespread use among children in and outside the home. This suggests that the source was probably not parental for most children. Parents often did not understand PE as well as their children and sometimes scolded them for not speaking their AL. Bickerton's scenario explains none of these facts.

The evidence also throws light on several other issues. It supports the *ca.* 1904 and 1910 dates for G3 emergence and the ranking of ethnic groups (in the order of Portuguese, Chinese, and Japanese) by demographic and sociolinguistic evidence. It also shows that Hawaiians patterned like the Portuguese.

6. Linguistic Evidence of Creole Formation and Nativization

The picture would not be complete without a discussion of structural innovations in the language of the locally born, as revealed by early pidgin/creole texts. This evidence will allow us to link demographic and sociolinguistic changes with the creole formation process.

One popular view places the formation of HCE in the 1880s or earlier and claims that much of the structure was diffused from a previously-existing pidgin or creole (Goodman 1985; Holm 1986, 1989; McWhorter 1994; Dillard 1995).⁹ This account was advanced as an alternative to Bickerton's scenario and rejects the possibility of locally-born involvement in creole formation. Foreign-born and Hawaiian adults would have created HCE by themselves since no significant population of G2 had emerged by the 1880s. The development of HCE would have resembled that of Solomons Pijin, described in section 2.

The facts do not support this proposal. Texts from 19th-century sources show that the early pidgin lacked nearly all the structures identified by Bickerton as creole innovations, including an integrated tense-modality-aspect (TMA) system with combinations of auxiliaries, an article system with specificity distinctions, and clausal complementizers (Roberts 1998). So there is no evidence that something like present-day HCE arose before the 20th century.

Structural differences between the PE of immigrants and the locally born were clearly evident by the 1930s; careful observers began to distinguish the

two as separate languages. In her study of Japanese substratal influence on HCE, Kaapu (1937:95-96) wrote:

That the pupils of the public schools now speak a language very different from the 'pidgin' used by the immigrant forebears of most of them, yet in varying degrees having remnants of the vocabulary and syntax of 'pidgin,' is quite evident to the careful observer.

Reinecke (1969:194), in his 1935 M. A. thesis, claimed that the pidgin of foreign-born adults "was the basis from which the children of the immigrants (and natives), educated in the English language schools, derived a form of English also commonly called 'pidgin English,' but more adequate and more refined than the makeshift speech of the immigrants". In an earlier letter to Elsie Wilcox (8 December 1932; University of Hawai'i, Hamilton Library, Hawaiian and Pacific Collection), Reinecke stated his belief that the two languages diverged around 1900:

[A]t approximately that date [1900] . . . the present dialect of what may be called 'the second generation' — and the third as well — and those influenced by them, then began to take form. . . [I]t is decidedly different from the English spoken by the immigrants who came to Hawaii as adults — the true Japanese-English, Portuguese-English, and other dialects — and also from the true 'pidgin' of the plantations; but . . . owes much to both the 'pidgin' and the national brogues of the Hawaiians, Japanese, and Portuguese (p. 1-2).

In Reinecke & Tokisama (1934), Reinecke described the structure of the "dialect" spoken by locally-born children; the article could almost pass as a description of present-day HCE. So it is clear that whatever changes distinguished HCE from the earlier pidgin were evident by the 1930s. But how long did it take for these innovations to develop?

In earlier work (Roberts 1998, 1999), I examined the first attestation dates of TMA markers *bin* (past tense), *go* (future/irrealis modality), *stei* (progressive/habitual aspect), *pau* (completive aspect), *wen* (past tense), and the combinations *bin stei* (past progressive), *go stei* (future progressive), *bin pau* (past completive), *go pau* (future completive), as well as *for* as a complementizer of verb phrases (or null-subject clauses) and finite clauses (containing nominative subjects). Also I divided the corpus into texts clearly attributed to children (or locally-born speakers born after 1890) and texts attributed to Hawaiians born before 1890, immigrants, or unidentified speakers.

These 11 features fell into two categories: items shared between foreign-born and locally-born speakers and items exclusive to the locally born. Table 7

below displays the results, in order of first attestation:

Table 7. Appearance of 11 HCE features in texts, 1881-1946

	Shared with locally born	Exclusive to locally born
<i>go</i>	1881	—
<i>bin</i>	1890	—
<i>for</i> (VP)	1891	—
<i>pau</i>	1896	—
<i>for</i> (clause)	—	1916
<i>stei</i>	—	1921
<i>bin stei</i>	—	1934
<i>wen</i>	—	1936
<i>go stei</i>	—	1937
<i>bin pau</i>	—	1946
<i>go pau</i>	—	1946

These data reveal the locally-born population as an important factor in the formation of HCE. Seven of the selected features were exclusive to the locally born and occur nowhere in contemporaneous texts of immigrant speakers. These include all the TMA auxiliary combinations and the clausal complementizer *for*.

The most striking pattern in Table 7 is the split between the locally born and the rest of the population between 1896 and 1916. Earlier features occurred in texts from both groups, though their frequencies were much lower outside the locally-born population. They were rare in foreign-born texts even in the 1930s and 1940s, when HCE was most influential. This suggests that they were probably not integrated into the overall structure of the foreign-born pidgin, even if they were first innovated by immigrant adults. As for *bin* and *for* (VP), locally-born texts attested these in very high frequencies. The Farrington quote in section 4 called the PE of locally-born children “the I-bin-go method of communication”, an apt name considering the abundance of *bin* in texts from the first two decades of this century. So even though *bin* and *for* (VP) were shared features, their distributions were significantly different in the two groups.

Later features occurred only in locally-born texts. These were most likely developed by the locally born. If one assumes a short time lag of about a decade between the innovation of these items and their first attestation in the texts, their development would span roughly 30 years from the mid-1900s to

the mid-1930s. This highlights the 1900 to 1910 decade as pivotal in the development of HCE, as demographic and sociolinguistic evidence has already shown. It also supports Reinecke's hypothesis that HCE began to diverge from the foreign-born pidgin after 1900.

Table 7 also shows that the process of structural elaboration proceeded at a much faster rate in HCE than in other documented creoles. According to early texts, the interval between the first attestations of TMA auxiliaries and combinations thereof spanned 88 years for Mauritian Creole, 80 years for Sranan, and 79 years for Antillean Creole French (Baker 1995; Corne 1995). The same developments took only 53 years in Hawai'i. If HCE had evolved at the same rate, auxiliary combinations would not have appeared until the 1960s. So HCE developed faster than other attested creoles, but not within a single generation as claimed by Bickerton. Yet if we disregard the early developmental period which did not involve the locally born (prior to the emergence of innovations by the locally born), only 18 years remained before the first cases of auxiliary combinations appeared. This allowed a single generation, born from 1905 onward, to witness by adulthood the expansion of HCE to something close to its modern state.

7. Synthesis

In this section I will assemble the preceding lines of evidence into a coherent picture of creole genesis. But first I would like to review and evaluate Bickerton's model of nativization as it pertains to the situation in Hawai'i.

The sketch at the outset of this paper highlighted three primary features of Bickerton's scenario: (1) the usefulness of ALs rapidly declined in Hawai'i, (2) G2 children, if exposed at all to their AL, did not acquire it, and (3) G2 children created the creole by acquiring their parents' pidgin. The bioprogram hypothesis rests on these assumptions. If children managed to acquire their AL and if it was useful for them in the community, then they were not at risk of growing up without a native language. If G2 children did not derive the creole from the pidgin of foreign-born adults, creole formation could not be linked to first language acquisition. Furthermore the primary empirical support for the bioprogram, the sharp contrast between the speech of immigrants and locally born (cf. Bickerton 1981, 1984a), would be undermined.

Bickerton has explicitly expressed these assumptions, though not without

contradictions. Below I contrast these statements with the evidence surveyed in this paper.

(1) *The usefulness of ALs rapidly declined.*

One is forced to conclude that a pidgin can creolize at any stage of its development, and that the period at which this step takes place will be decided, not by any internal development in the pidgin, but by the communicative needs of children, i.e., *whether the ancestral languages of their pidgin-speaking parents do or do not constitute adequate and feasible means of communication for them* [emphasis mine] (Bickerton 1977b:57).

One situation that stands between these poles [normal language acquisition and none at all] is surely that of the children of speakers of an unstable pidgin in a displaced community *where ancestral languages are of very limited utility* [my emphasis] (Bickerton 1977b:64).

[N]one of the available vernaculars would permit access to more than a tiny proportion of the community [emphasis mine], and . . . the cultures and communities with which those vernaculars were associated were now receding rapidly into the past (Bickerton 1981:5).

Sociolinguistic evidence shows that in the period preceding 1900, G2 children needed their ALs to communicate with members of their families and the wider ethnic community. ALs also formed the linguistic basis of social interaction between children of different ethnic groups. Therefore ALs were “feasible” and permitted access to much more than “a tiny proportion” of the population. The situation changed after 1900, or more particularly, after 1910. PE grew in importance and social relations no longer required multilingualism. But, due to the cessation of Portuguese and Chinese immigration, most G2 children born in the 1910s and 1920s were Japanese and Filipinos who largely resided on plantations. Because of segregation, children grew up in ethnically homogenous communities where they needed their AL to communicate with parents and other adults. So ALs remained indispensable to G2 children even in the 1920s, at which time the creole was already in existence. These speakers usually did not encounter PE until older siblings brought it into the home or they played with children from other ethnic groups. In some cases they knew no English until they enrolled in school at age 6 or 7. Hence the statements quoted above do not reflect the actual situation faced by G2 children.

(2) *G2 children, even if exposed to their AL, did not acquire it.*

Children are born of pidgin-speaking parents, *and in many cases fail to acquire active control of either parent’s tongue* [emphasis mine] — yet for

them the pidgin must be (a) too variable to acquire in the manner first languages are normally acquired, and (b) too inadequate to express a full range of topics with the degree of subtlety that being human demands (Bickerton 1976:2).

[Creoles] have arisen in colonies with a largely or exclusively immigrant population speaking a dozen or more mutually unintelligible languages, where the only means of communication common to all speakers was an immature pre-pidgin continuum. *Children born under such circumstances did not acquire the set of competing languages or any subset of that set; often they did not acquire even the rudiments of their parents' language(s)* [emphasis mine] (Bickerton 1984b:145).

Obviously they also heard an indefinite number of ancestral languages; *these, however, they ignored* [emphasis mine], precisely because the elaboration of the pidgin represented, to them, far less a task *than the learning of an ancestral language* [emphasis mine] and the subsequent transfer of features from that language to the nascent creole (Bickerton & Muysken 1988:302).

These claims are directly refuted by copious references in published sources and autobiographies to the widespread acquisition of ALs by G2 children. They did not “ignore” their AL and focus only on the PE they might have heard their parents occasionally use. Moreover children born before 1900 did learn “a subset of the competing languages”, since PE had not yet been established socially as the predominant medium. G2 children born after 1910, on the other hand, did not encounter “a dozen or more” or “an indefinite number” of ALs because of segregation.

Interestingly Bickerton has admitted that G2 children were actually bilingual in their ALs. Bickerton (1981:16) states: “All the older, locally-born subjects we interviewed spoke at least one other language besides HCE when they were children”. In striking contrast to the statements quoted above, he has more recently written: “Most of the first creole generation simultaneously acquired one or more of their ancestral languages” (1994:13).

This is a stunning contradiction and it remains unresolved. It poses two potential problems for Bickerton: (1) the possibility of substratal influence in creole formation is raised, and (2) children would have acquired the AL as their native language. Bickerton (1981, 1984a) has discussed the first issue extensively but the second has not yet been addressed.¹⁰

(3) *G2 children created the creole by acquiring the parents' pidgin.*

The child born of pidgin-speaking parents would seldom have had any other option than to learn that rudimentary language [emphasis mine], however

inadequate for human purposes it might be (Bickerton 1981:5).

[W]ell-formed input from mother [emphasis mine] cannot constitute even a necessary condition for children to acquire language; for, otherwise, creoles could not exist (Ibid., 139).

However few native speakers there were [in Suriname], these still had to have a native language of some kind — every other human being does. What does S[ingler] suppose that their language was? Does he suppose it simply reproduced the current pidgin? Or does he suppose that they learned natively one or another of the available African languages [i.e., ancestral languages] that would have given them access only to a fraction of the general population? *It is surely more plausible to suppose that they learned whatever variety of pidgin the adults spoke* [emphasis mine], but applied their faculté de langage to that pidgin to provide themselves with the full natural language that constitutes the birthright of every member of the species (Bickerton 1992: 312).

Unless Bickerton means that children learned pidgin by overhearing their parents talk to other adults, he implies that parents addressed children in “chaotic” pidgin. Never has he stated this explicitly; usually he refers to “input from pidgin-speaking adults”, which avoids the issue. In either case, he assumes that parents were the source of the pidgin which G2 children developed into the creole. Evidence from published sources and autobiographies indicate that the situation was the reverse: Older children learned PE at school or from neighborhood friends, they brought it into the home where younger siblings began to learn it, children then began to speak to their parents in PE while being addressed in the AL, and eventually parents began to speak with their children in both AL and PE.

According to Reinecke (1969:164), “[i]t is unusual for parents to use in their homes a language which they have acquired, usually imperfectly, in adult life”. This would have been even more unusual if they used it exclusively, in place of their AL. Japanese immigrants in particular valued the transmission of their native culture to the next generation and many enrolled their children in Japanese-language schools, which were popular in the 1910s and 1920s despite opposition by many whites (Reinecke 1969). The sentiment was well-expressed by Kazuo Kaneda, a student in William Smith’s sociology course at the University of Hawai‘i: “Unless Japanese parents use pure Japanese language, moral education cannot be given to the children at home” (H-52, A Sociological Study of the First and Second Generations of Japanese in Hawaii, 1926, p. 12).

This is why the general use of PE by parents was not a circumstance faced by G2 children, as noted by Adams, Jones, and Smith's informants. They indicated that Hawaiians and G3 Chinese and Portuguese learned PE natively from parents in the 1910s and 1920s but not G2 Japanese. G2 Chinese and Portuguese prior to the 1910s learned ALs as well.

These facts invalidate the three central assumptions of the bioprogram hypothesis.¹¹ The bioprogram fails because it places the entire burden of nativization and creole formation on G2 and affords no special role to G3. It demands that both occur in the first locally-born generation. Interestingly Bickerton once considered the alternative as a viable possibility. In Bickerton (1977a:333) he concluded that the TMA system of HCE developed over two generations, the fully integrated system not emerging until the generation of "later (monolingual) creole speakers" appeared. He hypothesized:

Let us suggest the following, rather tentatively: a creole grammar is built up over time. Some portions of its may be established in the pidgin period and it may not be necessary for the creole speakers to elaborate such portions. Other portions may be added by the earliest native speakers (or, where the pidgin does not creolize, by the second or third generation of non-native speakers). Other portions still may be added by the first generation of 'fully creole' speakers, i.e. the first generation that is monolingual in the creole. We would then be left with what looked, after the fact, like a complete language system which had sprung fully formed from the heads of the first 'fully creole' generation. The system looked like an integral innovation because all its parts fitted neatly, and because all its antecedent systems had disappeared, as a scaffolding disappears once a building is completed. In most parts of the world, those antecedent stages are gone past recovery; only in Hawaii have we been fortunate enough to be able to recover them (315-6).

It is unfortunate that he did not maintain this original hypothesis, as it receives considerable support from linguistic evidence. Table 7 shows that some early features (*go*, *bin*, *for*-VP) first appeared in texts in the 1880s and early 1890s and were later shared with the locally born. In fact, all early examples were attributed by writers to immigrants or adult Hawaiians, revealing them to be G1 innovations. By 1910 these features were stabilized as regular grammatical structures in the PE of the locally born. At this point the population was still mostly G2. In the 1910s and 1920s, as monolingual G3 speakers joined the locally-born population in large numbers, new structures (clausal *for*, *stei*, *wen*, auxiliary combinations) were added to the stabilized code. By the 1930s HCE was recognized as a distinct language.

I believe that a plausible scenario for the formation of HCE can be constructed from the evidence presented in this paper. According to the event hypothesis, the stabilization process would have begun by the mid-1890s after the occurrence of Event 2. Event 3 did not occur until the mid-1930s, so HCE had about 40 years to stabilize and develop as a separate language. In my opinion HCE stabilized from the mid-1890s to the 1910s, as G2 children shifted to PE for social interaction and members of G3 began to speak PE in place of their AL. Stabilization probably progressed faster in places such as Honolulu where PE was spoken intensively. In plantation communities segregation reduced the extent of contact between G2 Japanese and children of other ethnic groups.

The 1900 to 1910 decade saw four demographic changes in the locally-born group: (1) huge numbers of G2 Japanese entered the population, reducing the overall proportion of Chinese and Portuguese locally born, (2) G3 Portuguese and Chinese entered the population in significant numbers, (3) the first members of G3 Japanese emerged, and (4) locally-born Portuguese reached numerical parity with their G1 counterparts. Each of these developments probably accelerated the stabilization process and contributed to the shift from multilingualism to exclusive PE use. The rapid increase of G2 Japanese changed the school population from mostly Portuguese to more than half Japanese by the 1910s. The rest of the student body needed to learn either another AL to interact with the newcomers or to use the previously-existing PE more extensively. The latter presented a lesser burden for them. The acquisition (and general use) of PE by G2 Japanese would have automatically established PE as the major means of communication. Meanwhile substantial numbers of G3 Portuguese and Chinese were born, and their parents were bilingual speakers of PE who were probably more home in it than their AL. G3 children (not G2 as envisioned by Bickerton) were therefore in a position to learn PE natively in first language acquisition. Their emergence would have decreased the need for others to learn Portuguese and Chinese, as they probably did not know these languages well themselves. The same could be said of the G3 Japanese, though they were exceedingly rare in the population (as indicated in Tables 3 and 4). Finally G1 Portuguese were surpassed by the corresponding G2 and G3 groups by 1900. Since all members of the G1 group (except perhaps those who immigrated as babies) were fluent native speakers of Portuguese, their shrinking proportion led to fewer opportunities for G2 and G3 children to learn and use their AL.

The existence of linguistic and social norms by the 1910s shows that stabilization had taken place. Past tense marking and overt complementation of VPs were regular features of the PE spoken by locally-born children. Norms on the appropriateness of Standard English and PE in certain social situations were also established, as statements by contemporary writers indicated. At this time new structures were innovated, gradually elaborating HCE into the language it is today. This elaboration process raises a pertinent question: Was the “event” hypothesis, in its present application, correct in its predication that HCE would develop to be more radical (distinct from the superstrate) than Mauritian and other creoles?

Table 7 shows that HCE did develop a fairly radical basilect, containing auxiliary combinations and complementation of finite clauses with *for*. These are generally regarded as radical features. But lexically and syntactically, HCE is for the most part closer to English than Mauritian is to French. While agglutinated articles pervade the nominal lexicon of Mauritian (Grant 1995), HCE’s lexicon contains no systematic deviations from English (aside from Hawaiian vocabulary). Its syntax also incorporates English reflexives, relative pronouns, *wh*-interrogatives and other features which diverge structurally from lexifier forms in Mauritian and other creoles. Why did the event hypothesis fail to predict this? It does not take into account the fact that locally-born children were nearly all educated in Standard English (an opportunity slaves in other creole societies did not have), compelling them to learn the lexifier. This fact recommends a cautious approach to demographic evidence, which may be misleading unless one takes into account other mitigating factors.

But the very existence of a fairly radical basilect raises problems for Singler, who has claimed that greater exposure to the superstrate would lead to greater superstratal influence on the creole (Singler 1986, 1992). Locally-born children in Mauritius and elsewhere were generally not bilingual in the superstrate. So why did a basilect develop in Hawai‘i? One reason may be the strong stigma against Standard English which arose rather early in creole formation. This factor and the desire to speak a variety of PE different from G1 foreign born (identifying them as local citizens) may have motivated speakers to modify HCE in directions away from both the immigrant pidgin and the language of the dominant white culture. The basilect would serve both these purposes. At the same time an acrolect was necessary for involvement in white culture in terms of education, employment, and other concerns. These factors, in my opinion, could have led to an expansion of the stylistic dimen-

sions of HCE, creating a continuum of registers ranging from acrolectal English to the exclusive in-group basilect. Other interpretations are possible but this explanation is supported by known facts.

The three-generation pattern of language shift (with loss of the AL by G3) is not unique to HCE and has been observed in many other immigrant societies, particularly in 20th-century America, although Fishman (1985) notes that the outcome is far from inevitable; in certain cases the AL may persist longer or even supplant the competing language. When the AL persists longer than the initial generation (to G2 and beyond), substratal influence on the developing creole becomes possible. According to Corne (1994:296-297), G2 speakers typically are bilingual in both languages and may introduce into the creole substratal patterns from their AL when "there is a substantial measure of agreement. . . The third generation (G3) is also a kind of "hinge" generation. Its members have abandoned (lost) the ancestral L1, their L1 being the now emerged creole".

Corne (1994, 1995, in press) shows that Tayo, a French-lexifier creole spoken in St-Louis, New Caledonia, followed this developmental path. Tayo emerged between 1910 and 1920, when the locally-born population included both G2 and G3. Grammatical innovations which appeared in Tayo, such as causatives, relativization, and focalization, are analogous to those found in the original ALs of St-Louis, New Caledonia. Similarly, Jennings (1995) asserts that Cayennais, the French-lexifier creole of Cayenne, arose between 1700 and 1710, some 50 years after the first importation of slaves in 1660 (Jennings 1995). G2 children emerged significantly by the mid-1670s and "they would have acquired Fon [their AL] as their native language", though they were also exposed to the developing French-lexifier pidgin. The generation born in the 1690s (G3) then "would have begun to nativize this enriched variety which had formerly been no one's native language" (Jennings 1995:34-35).

The possibility of substratal influence in the formation of HCE is thus raised by this study. But until now no comprehensive study has assessed the extent of substratal influence in HCE. This will be the subject of a subsequent study by the author. But the data in this paper do offer some important clues as to where to look for substratal influence. Hawaiians and Portuguese were the earliest to shift between their AL and the creole, while Chinese were a little slower, and Japanese were the slowest. This implies that the former groups had the strongest influence on the creole before it stabilized, while the Japanese in particular had much less influence. What work that has been done on

the substrate languages of HCE seems to bear this out; Kaapu (1937) could find very little evidence of Japanese influence. Further work by the author will look into this matter in greater detail.

Notes

1. I presented an earlier version of this paper at the Society for Pidgin & Creole Linguistics Annual Meeting in London, England, on 28 June 1997. John Rickford, Genevieve Escure, John McWhorter, Elizabeth Traugott, Natalie Schilling-Estes, Will Leben, and Chris Corne have provided valuable comments on this paper. Any shortcomings are my own.
2. Although this paper was written for a creolist audience, many issues are relevant to general linguistics. Explicit definitions of terminology would thus assist nonspecialists. Unfortunately creolists have failed to remain consistent in the use of certain terms, so the definitions in this paper are my own. *Lexifier* refers to the language from which a contact language derives its lexicon. The language spoken natively by the ruling or upper class (which usually has most prestige in the community) is the *superstrate*. Often it is the same language as the lexifier; this is the case for the pidgin and creole English of Hawai'i. The languages spoken natively by foreign-born slaves or immigrants form the *substrate*. In this paper, these are also called *ancestral languages*. *Pidgins* and *creoles* are both used to refer to contact languages which arise to facilitate communication between different substrate groups and which have developed cohesive grammars distinct from the lexifier. The term *creole* is restricted to those which today include native speakers in the speech community. *Creole formation* is the process through which creole grammar comes into existence. I have adopted Winford's (1993) avoidance of *creolization* for the same process, since some writers use it to mean structural change and others use it to refer to *nativization*. This last term refers to the process in which a language acquires a native-speaking community over successive generations. The ancestral language is usually abandoned in the shift to the new language. *Jelling* is the process in which creoles stabilize grammatically as distinct systems. A person born in the locality where the contact language develops (in this case, Hawai'i) is *locally-born* while others are *foreign-born*.
3. Actually not all whites were native English-speakers. Of the 2,520 whites, 272 were German and 81 were French. This number does not affect the dating of Event 1.
4. There are some well-known problems with Hawai'i's census data, particularly the 1900 census which was the first done by the U. S. Census Bureau. In the 1900 census there was an apparent miscount in the number of Part-Hawaiians (many of whom were partially white) and whites (Nordyke 1989). This does not affect our data since Tables 1, 2, and 3 omit both Hawaiians and whites. Due to the 1898 annexation of Hawai'i, people born on the U. S. Mainland were now counted as locally-born in the totals. Since they were not locally-born by my definition, I have added their number to the foreign-born totals for the 1900, 1910, 1920, and 1930 census. The 1878, 1884, 1890, and 1896 census contain locally-born figures which match the figures derived by subtracting the foreign-born from the overall population.

5. Interestingly, 46% of foreign-born Portuguese arrivals before 1900 were counted as children by the Bureau of Immigration (Takaki 1983:124). This is because the Portuguese immigrated in large families (unlike the Chinese and Japanese). The significance of this is discussed in section 5.
6. The careful reader will note that 50 Portuguese were already locally-born in 1878 (Table 3). Prior to the 1878 start of massive Portuguese immigration, there were already over 400 residents of Portuguese ancestry, many of whom had previously worked in the whaling industry (1830s-1860s).
7. These are rough estimates, of course. The dates only suggest the approximate time when at least half the number reported in the following census were born. Since the population grew gradually year by year, this could have also occurred in 1892, 1894, or even 1895 for G2 Japanese.
8. Kaapu (1937:64, 90) similarly wrote: "The 'pidgin' used by the older generation is more distinguishable by races than the dialect spoken by children. . . Japanese children have first formed their habits of English idiom and syntax from the Non-Japanese playmates who are earlier fluent in the use of a dialect very little different from 'pidgin,' or from older brothers and sisters who got their first spoken English from that source, or from grown-ups who spoke only 'pidgin' ".
9. McWhorter (personal communication, June 2, 1998) has since withdrawn his support of the Goodman/Holm proposal.
10. Bickerton may argue that children still needed PE as their primary *lingua franca* with each other and therefore turned to a bioprogram to expand the language for their purposes. But this scenario places creole formation outside the realm of first language acquisition and into the world of social interaction. Without the pressure to acquire PE as an exclusive native language, it is difficult to see how a biological explanation is preferable over a social one.
11. Bickerton (1992:312-313) has claimed that biologically-driven creole formation occurs regardless of population size; hypothetically even a single child could create a creole when the right social forces exist to allow the language "to spread like wildfire through the general population". On this basis Bickerton could disregard all the evidence presented in this paper because a G2 child could have conceivably existed, unobserved, somewhere on the plantations of Hawai'i whose parents spoke only pidgin, thereby denying him or her their AL. But this claim is unfalsifiable and therefore lies outside critical inquiry; the burden of proof rests on Bickerton to demonstrate that creoles indeed can originate in individuals with a biological, non-social impetus.

References

- Adams, R. 1925. "English in the public schools". *Hawaii Educational Review* 13: 1-3.
- Anonymous. 1886. *Biennial report of the president of the Board of Education to the Legislative Assembly of 1886*. Honolulu: Hawaiian Gazette Press.
- Anonymous. 1896. *Thirty-third annual report of the Hawaiian Evangelical Association*. Honolulu: The Friend.

- Anonymous. 1897. *The Report of the Minister of Public Instruction*. Honolulu: Hawaiian Gazette Press.
- Arends, J. 1989. *Syntactic developments in Sranan*. University of Nijmegen dissertation.
- Arends, J. (ed.) 1995. *The early stages of creolization*. Amsterdam: John Benjamins.
- Baker, P. 1982. *The contribution of non-Francophone immigrants to the lexicon of Mauritian Creole*. University of London, dissertation.
- Baker, P. 1995. "Some developmental inferences from the historical studies of pidgins and creoles". In Arends, 1-24.
- Baker, P. (ed.) 1995. *From contact to creole and beyond*. London: University of Westminster Press.
- Baker, P. and Corne, C. 1986. "Universals, substrata, and the Indian Ocean creoles". *Substrata versus universals in creole genesis*, P. Muysken and N. Smith (eds), 163-183. Amsterdam: Benjamins.
- Bickerton, D. 1976. "Creole tense-aspect systems and universal grammar". Paper presented at Society for Caribbean Linguistics Meeting, Georgetown, Guyana.
- Bickerton, D. 1977a. *Change and variation in Hawaiian English, Vol. 2: Creole syntax*. University of Hawai'i: Social Science and Linguistics Institute.
- Bickerton, D. 1977b. "Pidginization and creolization: Language acquisition and language universals". *Pidgin and creole linguistics*, A. Valdman (ed.), 49-69. Bloomington: Indiana University Press.
- Bickerton, D. 1981. *Roots of language*. Ann Arbor: Karoma.
- Bickerton, D. 1984a. "The language bioprogram hypothesis". *The Behavioral and Brain Sciences* 7: 173-221.
- Bickerton, D. 1984b. "The language bioprogram hypothesis and second language acquisition". *Language universals and second language acquisition*, W.E. Rutherford (ed.), 141-165. Amsterdam: John Benjamins.
- Bickerton, D. 1992. "The sociohistorical matrix of creolization". *Journal of Pidgin and Creole Languages* 7: 307-318.
- Bickerton, D. 1994. "How to acquire language without positive evidence: What acquisitionists can learn from creoles". MS.
- Bickerton, D. and Muysken, P. 1988. "A dialog concerning the linguistic status of creole languages". *Linguistics: the Cambridge survey, Vol. II*, F.W. Newmeyer (ed.), 268-306. Cambridge: Cambridge University Press.
- Bickerton, D. and Odo, C. 1976. *Change and variation in Hawaiian English, Vol. 1: General phonology and pidgin syntax*. University of Hawai'i: Social Sciences and Linguistics Institute.
- Blount, J. 1893. *Papers relating to the mission of James H. Blount, United States Commissioner to the Hawaiian Islands, prepared by the U. S. Department of State*. Washington, D. C.: Government Printing Office.
- Breckons, E. 1915. "Cow-boy Tony". *Oahuan*, March 1915, 12-14.
- Brigance, W.N. 1938. "Good English in this American democracy". *Proceedings of the 18th Annual Hawaii Education Association Convention*, 29-34. Honolulu: HEA.
- Carden, G. and Stewart, W. 1988. "Binding theory, bioprogram, and creolization: Evidence from Haitian Creole". *Journal of Pidgin and Creole Languages* 3: 1-67.
- Chaudenson, R. 1992. *Des îles, des hommes, des langues*. Paris: L'Harmattan.

- Corne, C. 1994. "Relativization and thematization in Tayo and the implications for creole genesis". *Journal of Pidgin and Creole Languages* 9: 283-304.
- Corne, C. 1995. "Nana k nana, nana k napa: The paratactic and hypotactic relative clauses of Réunion Creole". *Journal of Pidgin and Creole Languages* 10: 57-76.
- Corne, C. in press. "Tayo causatives: The retention in a French-lexified contact-induced vernacular of transfers from New Caledonian Melanesian". *Te Reo* 40.
- Dillard, J. L. 1995. Letter to the Editor [28 July 1994]. *Journal of Pidgin and Creole Languages* 10: 219-20.
- Fishman, J. 1985. "Language maintenance and ethnicity". *The rise and fall of the ethnic revival: Perspectives on language and ethnicity*, J. Fishman, M.H. Gertner, E.G. Lowy, and W.G. Milán (eds), 57-76. Berlin: Mouton.
- Goodman, M. 1985. Review of Bickerton (1981). *International Journal of American Linguistics* 51: 109-37.
- Grant, A. 1995. "Article agglutination in Creole French: A wider perspective". In Baker, 149-176.
- Hall, R. 1966. *Pidgin and creole languages*. Ithaca: Cornell University Press.
- Holm, J. 1986. "Substrate diffusion". *Substrata versus universals in creole genesis*, P. Muysken and N. Smith (eds), 259-278. Amsterdam: John Benjamins.
- Holm, J. 1989. *Pidgins and creoles* (Vol. 2). Cambridge: Cambridge University Press.
- Hymes, D. (ed.) 1971. *Pidginization and creolization of languages*. Cambridge: Cambridge University Press.
- Jennings, W. 1995. "The first generations of a creole society: Cayenne 1660-1700". In Baker, 21-40.
- Jones, S. M. 1942. *Study of Kauai*. Honolulu: University of Hawai'i M.A. thesis.
- Jourdan, C. and Keesing, R. 1997. "From Fisin to Pijin: Creolization in process in the Solomon Islands". *Language in Society* 26: 401-418.
- Kaapu, M. K. 1937. *A study of the influence of Japanese syntax and idiom upon the spoken and written English of a group of ninth grade pupils*. Honolulu: University of Hawai'i M.A. thesis.
- Lefebvre, C. and Lumsden, J. 1989. "Les langues créoles et la théorie linguistique". *Canadian Journal of Linguistics* 34: 249-272.
- Lydgate, J. M. 1919. "Flu'-Inspecting on the Garden Isle". *Paradise of the Pacific* 31: 30-32.
- McKay, H. W. 1896-1901. *When I was two and twenty, January 1896-January 1901*. Hawai'i State Archives MS.
- McWhorter, J. 1994. Rejoinder to Derek Bickerton. *Journal of Pidgin and Creole Languages* 9: 79-93.
- Nordyke, E. 1989. *The peopling of Hawaii*. Honolulu: University of Hawai'i Press.
- Pinker, S. 1994. *The language instinct*. New York: W. Morrow & Co.
- Plag, I. 1995. "The emergence of *taki* as a complementizer in Sranan: On substrate influence, universals, and gradual creolization". In Arends, 114-18.
- Prescott, A. M. 1891. *Hawaii*. San Francisco: C. A. Murdock.
- Prescott, A.M. 1899. *Makapala-by-the-sea*. Honolulu: Hawaiian Gazette.
- Reinecke, J. 1969. *Language and dialect in Hawai'i: A sociolinguistic history to 1935*. Honolulu: University of Hawai'i Press.

- Reinecke, J. and Tokisama, A. 1934. "The English dialect of Hawaii". *American Speech* 9: 122-31.
- Roberts, S. J. 1998. "The genesis of Hawaiian Creole and diffusion". *Language* 74:1-39.
- Roberts, S.J. 1999. "The TMA system of Hawaiian Creole and diffusion". *Creole genesis, attitudes, and discourse: Studies celebrating Charlene J. Sato*, J. Rickford and S. Romaine (eds), 45-70. Amsterdam: John Benjamins.
- Samarin, W. 1993. "The creolization of pidgin morphophonology". Paper presented to the Society for Pidgin and Creole Linguistics, Amsterdam.
- Sankoff, G. 1977." Variability and explantation in language and culture: cliticization in New Guinea Tok Pisin". *Linguistics and anthropology*, M. Saville-Troike (ed.), 59-75. Washington, D. C.: Georgetown University Press.
- Schmitt, Robert. 1977. *Historical statistics of Hawaii*. Honolulu: University of Hawaii Press.
- Singler, J. V. 1986. Short note. *Journal of Pidgin and Creole Languages* 1: 141-145.
- Singler, J.V. 1990. "On the use of sociohistorical criteria in the comparison of creoles". *Linguistics* 28: 645-669.
- Singler, J. V. 1992. Nativization and pidgin/creole genesis: A reply to Bickerton. *Journal of Pidgin and Creole Languages* 7: 319-33.
- Singler, J. V. 1995. "The demographics of creole genesis in the Caribbean: A comparison of Martinique and Haiti". In Arends, 203-232.
- Sokolov, J. and Snow, C. 1994. "The changing role of negative evidence in theories of language development". *Input and interaction in language acquisition*, C. Gallaway and B.J. Richards (eds), 38-55. Cambridge: Cambridge University Press.
- Takaki, R. 1983. *Pau hana: Plantation life and labor in Hawaii, 1835-1920*. Honolulu: University of Hawai'i Press.
- Weimer, A. 1934. "Meeting the 'Pidgin' English problem in our schools: What can we do to improve spoken and written English?" Paper presented at the Secondary School Principal's Conference, Honolulu, March 19-22, 1934 (Unpublished volume of proceedings, Hamilton Library, University of Hawai'i at Manoa, 50-59).
- Winford, D. 1997. "Creole formation in the context of contact languages". *Journal of Pidgin and Creole Languages* 12:131-51.

The Status of Sango in Fact and Fiction

On the one-hundredth anniversary of its conception¹

William J. Samarin

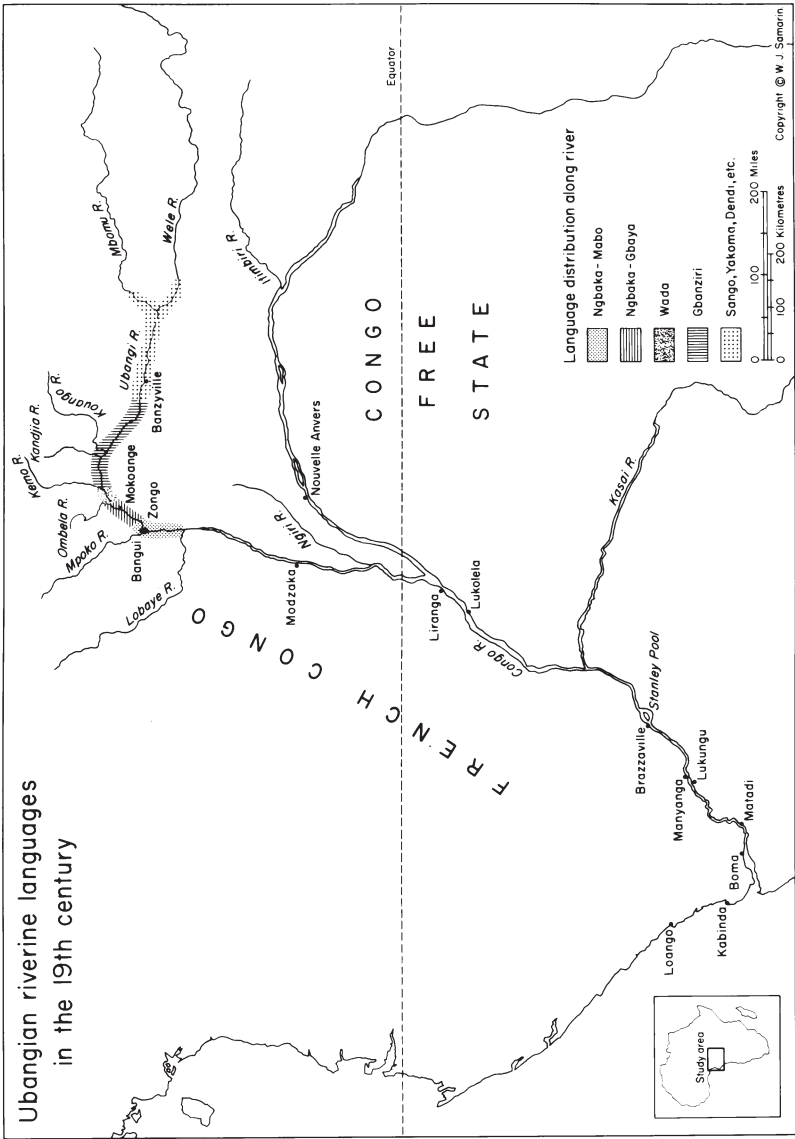
1. Introduction

Sango is one of the two official languages of the Central African Republic, the other being French. It acquired this role a few years ago, when André Kolingba was still president, a person whose native language is Yakoma, one of the several forms of Ngbandi, from which Sango is derived.² Before then, it was the nation's *langue nationale*, a role that was enshrined in the constitution at independence. Another fact, as revealed by the census of 1988, is that the percentage of Central Africans who know Sango is very high, but varying slightly from forty percent in the far east to ninety percent in some large towns or so-called *villes*; in the capital, as one might expect, almost everyone claims to know Sango; it would be embarrassing not to be able to do so.³ It is also a fact that in Bangui about forty-five percent of pre-school children (up to the age of about six) know no other language than Sango; they are therefore native speakers of this language (Samarin 1995).⁴

There are many other facts that can be cited without controversy. These, however, are not the subject of this work. Instead, I am concerned about some *alleged* facts: claims made about Sango that are contestable and that may indeed in some instances be false.

2. Questionable Historical Reconstructions

It is *not* a fact, for example, that Sango — by whatever other name or names it may have had at one time — was a *langue véhiculaire* (which is a synonym of



lingua franca) before the arrival of Europeans at the end of the nineteenth century, despite this having been proposed by a French *colon* long ago (Bruel 1918) and advocated by Diki-Kidiri (1977, 1979a, 1979b, 1981, 1982, 1985a, 1985b, 1985c, 1986, 1987a, 1987b, 1992). I have provided a great deal of information to discredit this claim, demonstrating that Sango arose as a new language (and in some ways a new *kind* of language), as the result of the immediate need that colonial interlopers had to communicate with the people already living along the banks of the Ubangi river. I refer not so much to the Europeans (Belgian, Scandinavian, Dutch, and French) but more to the Africans who were brought from different parts of the continent to help the whites in exploring, trading, and taking possession of land and whatever else they wanted. These Africans spoke not simply different languages but also *typologically* different languages: Niger-Congo (West Atlantic, Kru, Kwa, Mande, and Benue-Congo [i.e., Bantu, from the west, south, and east]; Afro-Asiatic (Hausa), as well as smatterings of European languages, the most important of which was English, probably one or more varieties of Pidgin English.

Therefore, at the end of the nineteenth century the main river basins of central Africa were a potpourri of languages. The sudden increase in linguistic diversity led, I believe, to the rise of Kituba and Lingala (Samarin 1986c, 1990), and Sango (e.g., Samarin 1982a, 1982b, 1984a, 1984b, 1984/1985, 1985, 1988, 1989a, 1989b, 1998a).⁵

The questionable history of Sango proposed by Diki-Kidiri is linked to his understanding of what the language once was, is today, and how it developed. Unfortunately, his description or characterization of the language through its alleged stages is as difficult to comprehend as his history, due to lapses in clarity of argumentation.

It is a fact, to be sure, that Sango is a language, containing all the necessary structures or systems of language: phonology, morphology, lexical units, syntax, semantics, etc. I have treated Sango like a real language — a real *African* language — the moment I began to learn it in 1952 (Samarin 1953, 1967a, 1967b, 1970). It is not, nor ever has been, anything else, except for the first few years of its existence when, according to the kind of linguistics I practice, it was in the *jargon stage*, and then an *incipient pidgin* (which is nothing more than a pidgin in the process of coming into existence). It then evolved into a *stabilized pidgin*, possibly as early as around 1910, and almost certainly by the 1920s. (At that earlier date Father Calloc'h was already working on his grammar and dictionary [Calloc'h 1911].) But for most of its

history, the French in their colony did not respect Sango, considering it even inferior to the preliterate African *dialectes*, as they frequently called all African languages. Since they never, to my knowledge, became competent in Sango or, except for a few missionaries, any African language, this must have been a prejudice they acquired from one another.

There has arisen some doubt, however, as to what *kind* of language Sango is. As the first person to study the language scientifically, I judged it to be a *pidgin* and called it that in one of my first publications on the language (Samarin 1958); earlier, I had only said that Sango was “simplified by the loss of most of its morphology and of the bulk of the original vocabulary” (Samarin 1955:261-62). After studying Sango for forty-five years I have discovered no reason to change my mind: Sango *is* a pidgin in origin. (The fact that it is the primary language of several thousand people now makes it also a *creole*, according to one of the well-known uses of this word.⁶ Diki-Kidiri, however, is of a different opinion; he does not consider Sango a pidgin, although to the best of my knowledge he has not in print said, “Le sango n’est pas un pidgin”, or words to that effect. Indeed, no one writing in French appears to have made such an assertion.⁷ On the contrary, according to French writers, Sango is today simply a language which was at one time a *langue véhiculaire*. An example is Gabriel Manessy, who refers to Sango as a vehicular language (1977:133; see also 1990:7, 1995:64) and a “parler véhiculaire” (1995:60fn), not as a pidgin (also Bouquiaux 1969).⁸

But let us review Diki-Kidiri’s analysis. (The following, including commentary, is in my own words; for his, see Samarin 1998a). To make the linguistic history of Sango easier to follow I assign numbers to the different stages of its putative evolution.

Stage 1.

Somewhere along the banks of the Mbomu river, a language was once spoken, slightly different, but only in the usual ways, from other varieties in the Ngbandi group. It probably did not have a name, because language names in many cases came to Africa, as in other parts of the world, with colonization. The people who spoke this local variety of Ngbandi also did not have a so-called racial or tribal name, identifying themselves by a clan or lineage name.

Stage 2.

At some point in the nineteenth century, this language came to be used by its people first in dealing with the co-territorial Ngbugu and Nzakara. This did

not mean that it was a vehicular language, however, because these two other groups did not use the language between themselves.⁹ If that had happened, the language would have become a lingua franca even as English is the world's lingua franca today.

Stage 3.

However, *because this language was used between different people, it changed slightly*, but no more than languages normally do when they come into contact. In other words, the most noticeable difference at first was probably in vocabulary, just as French words enter Arabic speech of people living in France. At this stage in its history the language was really the same as it was in Stage 1.

Stage 4.

The people who spoke this language were such successful traders along the Mbomu river that they descended it and entered the Ubangi, which flows westward, travelling about 800 km or more, “at least as far as Bangui” (Diki-Kidiri, personal communication, November 1996; cf. Diki-Kidiri 1981:35).¹⁰ This accomplishment would have been heroic, because these Central African Phonecians defied the challenge of the *Chutes Hannssens*, the rapids just north of Ouango. Moreover, they must have been so numerous, powerful, and skillful that they were able to intimidate their Ngbandi “cousins” (who quite early acquired from whites the names Sango and Yakoma) living along the banks of the Ubangi river. This account alleges that this people had dominated the trade in ivory, slaves, and iron in this river basin and traded with Sultan Bangassou on the banks of the Mbomu.¹¹ What is more surprising in this account of pre-colonial central Africa is the allegation that in paddling down the Ubangi river, these adventurous traders who had travelled right by their “cousins” not only traded further downriver with the Gbanziri-Buraka, the Ngbaka, and the Monzombo (people living just up- and downriver of Bangui) but even with the Banda who lived inland on both sides of the river, who were in general prevented from trading directly with those along the river. The “vehicular Dendi” designated as the trade language these people used is indeed called “langue de marché et d’échange” (Diki-Kidiri 1986:86). Crucial information about this trade is missing in Diki-Kidiri’s version of Sango’s history. A number of facts would help: for example, the frequency of the traders’ visits, the length of their stay on the Ubangi before returning to the Mbomu, whether or not they set up villages, what kinds of contracts they drew

up with local headmen, and the locations of the alleged markets. In other words, nothing is actually known of this particular alleged trade, which appears to be more a matter of oral tradition than scientific history.

Equally faulty is Morrill's history of Sango. While agreeing with Diki-Kidiri that there was a "vehicular variety of Ngbandi" (eshewing the term Dendi) in the precolonial period (Morrill 1997:166, and all other pages in this paragraph), he does not accept all his arguments and has his own, which I now number.¹² (1) "... Yakoma iron production placed them at the center of an extensive and well developed trade network." They in fact forged "the first currency of the region" of iron, "which were traded up and down the river and at a considerable distance into the interior". Their trade networks "fed Yakoma iron production", "carried Yakoma goods beyond the Ngbandi-speaking region", and "spread their language" (166); "before the arrival of Europeans into the region there existed vast kingdoms and complex trading networks" (50). (2) But the simplification of Ngbandi was due to slaves who "would have spoken a distinct foreigner variety" (166, 353). It was this form of Ngbandi that "contributed to the formation of a vehicular variety used in the Banda- and Nzakara-speaking region in the precolonial period" (167). The presence of a "considerable" number of slaves is assumed because "a great deal of labor" would have been needed for the "[l]arge-scale iron production" among the Yakoma (166, 353). (It would appear that although Morrill accepts Diki-Kidiri's vehicular Ngbandi [by whatever name], he imagines it an altered language, which Diki-Kidiri does not.)

This reconstruction of production and trade in the Ubangi river basin, itself the basis for a hypothesis of a linguistic nature, does not survive critical examination. A well-researched doctoral thesis (Zigba 1995) written by a Central African who learned Ngbandi for his field work and was directed by an anthropologist who has devoted his whole career to the study of this area supports my conclusions. Here are gleanings from his findings: (1) iron ore at shallow depths could be found here and there for "exploitation artisanale" (and in my research I learned of several places at the bend of the Ubangi river where iron was worked); (2) industrial exploitation, which would have required digging to greater depths, has never been documented; (3) it is doubtful that there were any specialists (e.g., a "maître-fondeur") in the villages; (4) the units of production were based on kinship, alliances, and friendship on a local basis; (5) the riverine populations had absolutely no reserves of slaves, and not

a single source mentions a Ngbandi campaign to acquire captives; (6) trade always took place with one's immediate neighbors, no one going "d'un bout à l'autre du circuit d'échange (1995:147; 150, 163, 169, 225, 246).

As for currency, Morrill has failed to consider the dynamics of colonization on the local economy and the relations between all the indigenous peoples. Yakoma-forged iron was not used (or not as extensively) as money in the west until beads had lost their value (i.e., because of inflation). In 1891–1892 the French commander Victor Liotard, in carefully noting his expenditures, repeatedly cites beads. Iron is not mentioned until the expedition was in the east, among the Yakoma. There, after the French had set up a market, the Gbanziri and Sango, Liotard notes in his journal, were able to buy *kinja*, the local iron money.¹³ (Liotard also mentions that Bangui became the haven to which Yakoma came when fleeing those they had mistreated while serving under the French).¹⁴ Thus the picture Morrill limns of a center of iron production and forging, made possible by mining and lumbering activities, and undertaken by a large number of slaves, appears essentially a romantic neo-Ngbandi reconstruction if not historical revisionism.

Stage 5.

It is understandable that if this trade by these allegedly powerful people from the Mbomu river went on for a century or more, the local people would have learned some of their words. This much of language learning is what would be expected. The invaders, however, would probably not have done so (or done so to a significantly lesser degree), because, as is well known, people in a position of power do not commonly borrow more than a few words from their clients endowed with a role of less power. However, according to Didi-Kidiri, *by this time the language of the river traders had become a lingua franca*. This can only mean that it was not used only by A (identified as Dendi by Diki-Kidiri), the Mbomu traders as "*langue de contact*", but also between speakers of B and C, who had each learned it from A or had learned it from each other. Naturally, when the Gbanziri spoke to the Banda in A, it would not have been a perfect representation of it. But we have no evidence of what changes had taken place in the language. We cannot even guess. They would probably not have been really significant ones, because all these languages are related, being members of Joseph Greenberg's Eastern (now more often referred to as Ubangian) group. Moreover, change would have presumably been going on very slowly.

Stage 6.

The greatest changes took place when the language began to be used by the colonizers. Despite the impression lent by books and articles written by whites of the colonial period, European colonizers were few in number by comparison to all the Africans they brought with them into the Ubangi river basin. Nevertheless, according to Diki-Kidiri's account, this was when the lingua franca, allegedly first known as Dendi and then as Sango, underwent more changes. The changes, according to some, were deliberate on the part of the Sango, who simplified their language for the interlopers (Bouquiaux et al. 1978:21).¹⁵ But, according to Diki-Kidiri, the changes were not drastic ones.¹⁶ Of course, some of its vocabulary was changed (although no examples are provided by Diki-Kidiri) as well the syntax to a minor extent.

Fact or fiction? What can we believe in what has been only sketched out for us in different publications? Let us not here argue about any of the details of its history. *The only thing we would like to establish is whether or not Sango is a pidgin.*

3. The Real Sango

There are at least two reasons for affirming that Sango is a pidgin.¹⁷ The linguistic changes that have separated Sango from its once-closely related Ngbandi dialects are (1) drastic, and (2) they developed very quickly. The following brief exposés must suffice for the present purposes.

3.1. Rapid change

Rapid divergence

Although all languages change in ways that make them more and more unlike the ones that preceded them, they change slowly. Of course, the rate of change is not so steady that we can arrive at an equivalent to carbon dating: languages change at different rates, and they have in many instances probably changed at different rates at different periods of their histories. The rate of change is also variable with respect to different systems: phonology, morphology, etc. All things considered, nonetheless, languages do not usually become vastly different in one or even two generations. The youngest child of all speech communities always speaks the same language (although not necessarily in

the same way) as the eldest. These facts have become commonplace, for many of the principles of linguistic change were long ago established.

Sango, however, like other pidgins in existence today, came into being very quickly.¹⁸ They must be cited as exceptions when generalizations are made about normal language change. Certainly they are not “abnormal” languages, nor are their changes “abnormal”; the only fact affirmed at this point is that these languages — at least those whose histories can be traced to the last century — diverged from their lexifiers very quickly, although the degree to which they are different from other languages is a matter of current debate and discussion.

It is therefore of utmost importance to recognize the rapidity of Sango’s emergence (within a short span of the lives of Central Africans!), to understand its significance, and to reflect on its consequences. In about 1890 along the Ubangi river there were native speakers of some variety of Ngbandi, who had spoken with others whose speech was slightly different from their own (i.e., other dialects of Ngbandi), and who had also heard about — if not encountered — others whose languages were entirely different (such Nzakara, Zande, and varieties of Banda). And right before their noses, so to speak, a new new *yángá* (as ‘mouth’ and ‘language’ came to be called in the pidgin) came into existence before many of them had died. They most likely knew nothing of pidginization and pidgins. The argument here is that the *fact* of rapid divergence in Sango’s history obligates a linguist to consider pidginization in its genesis.¹⁹

Drastic divergence

Pidginization is something that happens to language in different kinds of circumstances, one of them being the learning of another language, but perhaps usually only in adulthood (Samarin 1971).²⁰ All kinds of mistakes are made in the early stages of language learning, even when the circumstances are ideal for acquisition. A hundred years ago the circumstances along the Ubangi river were far from ideal. Moreover, it cannot be assumed that many of the Africans brought into the area by whites were even trying to learn a local language; i.e., aiming to talk it more or less like the native speakers. Verbal communication being their immediate need and goal, it can be justly assumed that they used whatever lexical resources and rules that seemed to work. (Even if they had the desire [to learn another language], they did not have the opportunity to do so, since they were moved about so frequently and many

knew that they would be returning home after their contracts expired.) Wasn't it inevitable, therefore, that the *jargon* or *baragouin* or *charabaria* or *patois* (in French) resembled only superficially the language of the upper Ubangi basin?²¹ However, after an infinite number of extemporizations, many of them repeated and imitated by others, *intention would have replaced chance*. Out of this process came into being a linguistic *product*: the new language, a pidgin: it became something that a person learned from others, not created personally.

The pidgin's drastic divergence from Ngbandi was inevitable, although the degree and nature of that divergence depended on many linguistic and nonlinguistic factors that constituted the circumstances. The Ngbandi experienced the differences and had to overcome them, a process that was *their* contribution to the creation of the new language.

Before proceeding, we should recognize the opposite view: that Sango is not nor ever has been a pidgin. As noted above, Diki-Kidiri has never made this assertion, but he appears to hold it; Morrill, on the other hand, declares it categorically: to classify Sango as a pidgin is an error, because "The *sine qua non* of what it means for a language to be a pidgin [elsewhere creoles are added] must be genetic discontinuity with any one natural language"; there has to be "a break in normal linguistic transmission" (1997:3, cf 10, 33, 35, 352). Therefore, to prove that Sango is not a pidgin Morrill has only to demonstrate that it is not a new language but simply another "distinct variety of Ngbandi." The evidence is never presented, but the claim is made again and again (1997:3, 34, 35; and all pagination in this paragraph). Without any review of the debate over whether or not pidgins can be or should be assigned a genetic classification, Morrill asserts that "it is an integral aspect of pidginization that the result of this process [i.e., a pidgin] is a new linguistic creation which may not be placed onto any existing family tree" (352). On the contrary, it can also be argued that Sango can be a pidgin *as well as* be closer to Ngbandi than to any other language.²² In any case, while he recognizes that changes have indeed taken place in Sango (350-51), he maintains that they have not severed it from Ngbandi's tree — yet this thesis is never defended. Nonetheless, Morrill proposes "the recognition of vehicular languages as a distinctive linguistic category" (26), to which Sango is assigned, because "while it can be attached to a specific genetic stock, it does not represent 'normal' transmission" (35, cf 352, 353).

This proposal might appear somewhat eccentric in light of most pidgin and creole scholarship, but the explanation is found in Morrill's having

adopted for his theoretical framework the work of Louis-Jean Calvet (1992), whose underlying notion is said to be, in Morrill's words, "that changes in the social function of a language may induce changes in grammatical form" (1997:350). However, Morrill ignores some of Calvet's opinions that could very well apply to Sango. The latter, for example, identifies a pidgin as a language *created* [italics in his text] to serve as a lingua franca, used by people who do not share the same vernacular ("langue grégaire"), "le plus souvent *composites* ..." A pidgin is therefore "*in vivo*, ce que l'espéranto est *in vitro*, un certain type de réponse à un besoin de communication dans des conditions particulières", citing Munukutuba (more commonly known among linguists as Kituba) as an example (1992:18; also 23; also 1986:299). A pidgin is therefore no one's first language (1992:19).²³ The creative process itself is described in an earlier work (not cited by Morrill): pidginization is the "création d'une langue à partir de plusieurs langues naturelles" (Calvet 1974:114) because of the "refus pur et simple d'un groupe de parler la langue des autres groupes avec lesquels il entretient des relations commerciales", citing Pidgin English [sic], based on Chinese syntax and English lexicon (1974:113). If this were applied to Sango, it would mean that the Ngbandi refused to speak to the foreigners in their own language. It would also mean that because Sango's syntax and vocabulary are basically those of Ngbandi, Sango is not a created vehicular language — therefore, not a pidgin. Of course, one is not obliged to adopt Calvet's premise. Reasoning differently and historically, one arrives at a different conclusion.

That Sango is "genetically related in the normal sense of the term" to Ngbandi is Morrill's conclusion, based, he says, on "correspondences in all areas of the grammar." He believes that "It is [the] overwhelming similarity of basic vocabulary that provides the most conclusive linguistic evidence that Sango and Ngbandi both stem from the same linguistic source" (1997:217). (Here he is saying that the two languages are sibling descendants of an earlier language.) This conclusion is based on a study similar in purpose to my etymological analysis of Sango's vocabulary, using M. Swadesh's one-hundred-word list (Samarin 1961).²⁴ Because his method was very different, his conclusion is very different from mine. In any case, of the one-hundred "core vocabulary items" he finds only the word for 'green' missing in both Ngbandi and Sango and two others that were, he claims, borrowed before 1900 (Morrill 1997:214). This is an important discovery for him (since I claimed that the percentage of retention was lower), because if Sango were a pidgin, he

believes, “one would expect to find from its very inception a heavily mixed vocabulary drawing lexical items from the numerous substrate languages present at that time” (1997:214).

The following observations constitute a brief critique:

(1) This is a self-imposed constraint: Sango’s vocabulary does not have to be heavily borrowed for it to be a pidgin. A study of Cameroon Pidgin using the two-hundred-word list confirms my hypothesis in finding that the borrowings in this pidgin constitute no more of its basic vocabulary than the borrowings in German, French, and Spanish (Gilman 1979). Therefore, that most of Sango’s vocabulary is clearly derived from Ngbandi is irrelevant to its being or not being a pidgin.

(2) The core or basic vocabulary used in this study are known to be the ones *last* borrowed in languages (Huttar 1996). Indeed, everyone recognizes that a pidgin has a *lexifier* language. In Sango’s case, it would be Ngbandi.

(3) Morrill’s findings and mine are not comparable, most importantly for spoken Sango. While I elicited words from rural speakers, he used published dictionaries exclusively, some of them influenced more or less by sources of Ngbandi ethnicity. (My study cautioned against a naive use of existing dictionaries [Samarin 1961:17].)²⁵ Since not all these dictionaries were based on words heard in actual use, some of the entries might not represent Sango in its popular form. For example, although I was not able in the 1950s to elicit words for *tongue* and *ashes*, I knew that they were in the Bible, having been elicited (I have reason to believe) from Yakoma-speaking assistants. Another reason for having found fewer Sango words to complete the list than Morrill did was that I was rigorous: for example, I did not accept *bé* for *heart*, knowing that this was a westerner’s metaphorical calque — the word really means ‘liver’ (as it does in Ngbandi [Lekens 1958:47a]).

Moreover, in the long view, Sango, as I know it, is different from every variety of Ngbandi in several profound ways. In the following overview the two languages are compared with respect to phonology, morphology, syntax, and lexicon.²⁶

3.2. Linguistic characteristics

Phonology

The phonology of Ngbandi, unlike other systems of the language, was largely retained. The phonologies of the two languages are, giving allowance for interpretation, virtually congruent. This fact does not, however, constitute evidence against Sango's being a pidgin; it is rather evidence that many generalizations about pidgins have been simplistic, although it has wisely been pointed out that in pidginization all systems are not necessarily affected, and those that are are not affected to the same degree (Thomason 1981). Linguistic and nonlinguistic facts, as in every case of language contact, can, if data are available, explain the results.

The explanation for Sango's full phonemic inventory — unlike that of, for example, Fanakalo — lies partly in the phonological systems of the languages whose speakers participated in the creation of Sango, many of whom were West African, and the languages spoken in Haut-Oubangui, some of whose speakers were quickly exposed to this new language. Finally, it lies in the fact that the Ngbandi of this territory (the Yakoma and Sango) repossessed their language. (Evidence for this argument will be presented in due course.) Simplification is most readily seen in words that came from other languages; e.g. *nsinga* 'cord, telegraph wire' (in Lingala, but probably from an earlier source) > *singà* 'telegraph, telephone'.

Simplification in Sango is seen primarily in the use of its phonological inventory. (1) The number of words with nasalized vowels is very small in Sango. (2) The tonal characteristics of words have been simplified by reduction in the use of contour tones and in the sequencing of tones in words: e.g., mid-low has been replaced by high-low.²⁷ (3) Co-occurrence of vowels has been simplified by vowel harmony: i.e., mid vowels in a single word are either tense or lax, not both. (4) Tone changes in Ngbandi, whose rules have yet to be written, appear commonly: e.g., *álà* > *lâ* as in *tèlé lâ* 'their bodies' (body 3P) (Toronzoni [henceforth T] 1989:527), *nwó lâ* 'overcome them' but *nwô í* 'overcome you (pl.)' (T382, 383), *nà* > *nâ* 'with'.²⁸ In Sango, on the other hand, only the forms with level tones are used. My explanation for the survival of Ngbandi phonology in Sango is that the Ngbandi reconstituted the originally 'mispronounced' words as they themselves spread the pidgin in the territory, something similar in consequence, but not in process, to decreolization. The Yakoma in Bangui continue to have an influence on Sango.²⁹

One aspect of Ngbandi's phonology that was regularized in pidginization was the elimination of the use of floating tones in all but a very few words. Thus, in Sango (at least in the 1950s) the vowel of the pronoun *lò* was pronounced long, with rising pitch, before the word *mvèni*: *lò mvèni* 'he himself/she herself'. In other cases, it would appear that what seems like a floating tone in Ngbandi is only the result of, for example, assimilation: thus, in a predicate, where the verb occurs with the subject marker *à*- one frequently finds in Lekens' dictionary a tone mark: e.g., *`sùmbù* (1958:65b). Here too, things became simpler (i.e., more regular) in Sango: that is, until very recently, when Ngbandi-like contractions were noted in the tape-recorded speech of young people in Bangui.

Morphology

The Ngbandi, however, could not reclaim the lingua franca to make it like their own language grammatically. Although they may have continued to pronounce words as if they were their own, speaking Sango with a Ngbandi accent, they had to follow the grammatical rules of the new language.

Ngbandi's morphological grammar has virtually disappeared, making the grammar of Sango drastically simple when compared with the grammars of all the other Ubangian languages.³⁰ This is seen in the verbal system, in nominal derivation, and in possession. The details of the drastic changes in the grammar cannot be summarized in the space permitted in this chapter, but they can easily be seen in comparing what Benjamin Lekens (1955, 1958), Pascal Boyeldieu (1982), Ngama-Nzombio Tra Ndele Toronzoni (1989), and Kola Kamanda (1989) have written on the language and what has been published on Sango.

The verb. The Ngbandi verbal system is especially complex. Boyeldieu describes three verbal categories: *positif*, *inachevé*, *volitif*. These, in turn, are complicated by differences between singular and plural. Only the first is illustrated here. (1) The Ngbandi could no longer use their language's complex verbal system to express tense, mood, and aspect. The following is only suggestive of what happened. In Ngbandi two aspects, the perfective (accomplished action) and imperfective (action not yet accomplished), are expressed entirely by tone. This is complicated by the fact that all verbs do not behave alike. For example, irrealis (as in what is translated as future in English and French) is realized in the second person singular with a high tone on the pronoun but low, mid, or high on a monosyllabic verb, depending arbitrarily

on its classification. Therefore both the tones of the pronouns and of the verbs may change: e.g., *mbi hu lò* 'I saw him/her' vs. *lò hú yé* 'he/she saw thing(s)' (Lekens 1958:301a). With a preposed *ndó* (possibly a prefix, possibly a free form), whose tone will vary, the habitual or "intemporel" is expressed (T238). Sango has none of this, except for frozen forms like *ádè* (SUBJ+remain) 'it remains, not yet', where the prefixal subject-marker carries high instead of low tone: e.g., *ádè lò ga apè* ([he/she] remains 3S come NEG) 'he/she hasn't come yet'.³¹ The habitual marker was lost completely, except in the Protestant word for 'love' as both noun and verb: *ndòyé* (pronounced somewhat in that manner).³² Tense, mood, and aspect are in Sango expressed lexically, except for the use of the copula for irrealis. But the copula itself and therefore this use are foreign: innovations, in other words (Samarin 1986c). (2) They did not have available to them Ngbandi's means to produce new verbs with suffixes: e.g., the iterative in its various forms (e.g., *-là*, *-ngàrà* (T126), and the reciprocal *-ngbi*, *-ma* to derive verbs from ideophones (T235), and *-kà* (T145). Indeed, it is possible that synchronically some or all of these suffixes were no longer productive even in Ngbandi. In any case, in Sango, although a few words with these strings occur in the lexicon, there is no relationship between them and a base form. Thus, although some people may use the word *lèkèrè*, they do not distinguish it from *lèkè*. The two words mean nothing more than 'to repair', not with an iterative sense.³³

The noun. In Ngbandi, nouns can be derived from ideophones by reduplication with tone modification (T128-29, 176) and from verbs (a) by means of tone change (*nzì* 'to steal' > *nzi* 'theft') (T132-33, 168, 170, 175) and (b) by suffixation with *-ngo* (T172-174). Nouns are also derived by compounding (producing *substantifs composés*) (T177-89). In Sango only the nominalization of verbs is unrestrictedly productive: e.g., *gi* 'to look for' > *gingo* 'search'. While compounding has produced some words, the number is not as great as in Ngbandi. Indeed, the analysis of this aspect of Sango's grammar remains problematic. For example, since *yángá-dà* and *yángá tí dà* (mouth of house) are synonymous, should the latter be considered a compound instead of a noun phrase? Toronzoni opts for the former (T187), but I (and Roulon 1972) for the latter.³⁴

The adjective. The nineteen Ngbandi monomorphemic adjectives (T219) are increased in number by derivation from verbs in two ways, reduplication of monosyllabic verbs (with tone change) and tone change with disyllabic verbs: e.g., *sé* 'to be bitter' > *sèsé* 'bitter' (T125); *vùlù* 'to become ripe' > *vùlu*

‘ripe’ (T127). Although these two words occur in Sango, other pairs do not — further evidence that the process is not productive in Sango.

The numerals. Ngbandi has two sets of numerals: one for counting and another in attribution (T293). The latter are shortened forms of the other. Numeration in Sango, as one might expect in a pidgin, has only one paradigm, this consisting of the independent forms.

Syntax

Negation. Sentences in Ngbandi are negated with a preposed and a postposed clitic: *má* (or *tá*)... *ma*, as in standard French *ne ... pas* (T462, 529) or a simple postposed *ka* (T467). Sango uses only a postposed *apè* (*pepè* in published material).³⁵

Ngbandi:

- (1) *dema tá á-mu yé ma*
 groaning NEG SUB-grab thing NEG
 ‘Groaning doesn’t do any good’

Sango:

- (2) *tòtò à-sàrà yé apè*
 to cry SUB-do thing NEG
 ‘Crying doesn’t do any good’

Pluralization. In Ngbandi plurality with substantives (and in noun phrases) is signalled by the use of the prefix *á-*, in a few cases with a replative form of the noun: *ya* ‘wife’, *áwó* ‘wives’; *zò* ‘person’, *ázì* (or just *zì*) ‘persons’ (T210, 365). Sango continues to use the prefix, but makes no use of replative forms. With Ngbandi verbs, plurality, whether in the subject or object of a verb, is encoded with replatives (T235, 365, 514, 515): e.g., *sa* replaces *bi* when one person throws more than one spear and several people throw spears (T514). Even adjectives are differentiated for number: *kota* in the singular and *kokone* in the plural for ‘large’ (T365). All of these restrictions have been lost in Sango.

Possession. As in other Ubangian languages, Ngbandi distinguishes between alienable from inalienable nouns in possession. With the latter, possession is expressed simply by juxtaposition of nouns or nouns and pronouns, in some cases with obligatory tone change, as in the following: e.g., *tère mbi* ‘my body’ (body 1S), *téré lò* ‘his/her body’ (Lekens 1958:319a). With a semantically defined class of nouns, plus nominalized verbs, a difference in meaning is made with the use and non-use of the preposition: e.g., *dà Nzàpa* ‘church’

(house God) vs. *dà tí Pídà* ‘Pida’s house’ (T363). With the latter either *tí* or *te* are used, which are selected according to whether the determining noun is animate or not. In Sango all of these restrictions have been lost and *tí* is the only preposition used: e.g., *dà tí Nzàpa* ‘church.’

Native speakers of Ngbandi undoubtedly realized that saying *lì tí mbi* for ‘my head’ in the foreign (not *foreigner*) talk was not what one would ever say in their language: it would have been *lì mbi*. Put quite simply: the preposition could not be used with inalienable objects. In the last thirty years or so, however, the use of the preposition has changed considerably in Sango (Samarin 1993, 1995, 1997). One of the new realizations is the reduction of the preposition to a high tone, which is applied to the final vowel of the preceding word: *dà tí mbi* > *dà R mbi* (house of 1S) ‘my house’ (R = rising tone on preceding vowel replacing preposition *tí* ‘of’). In Ngbandi, however, the final tone depends on the phonological environment: *(ya) nga mbi* (belly mouth 1S) ‘my mouth’ vs. *(ya) ngá lò* ‘his mouth’ (Lekens 1958:651a).

Relativization. In Ngbandi two demonstratives are used in relative constructions, *ko* and *mε*, according to certain constraints that are not discussed here (T531). Neither of these has survived in Sango, which uses *só*, now a demonstrative but in Ngbandi, apparently, an adverb meaning ‘there, thus, in this manner.’³⁶ (2) The relative clause has been transformed and reduced to the use of an all-purpose *só*.

Ngbandi:

- (3) *zò mē mbì hu lò bíri làá*
 person REL 1S see 3S yesterday TOP
 ‘that’s the person I saw yesterday’ (Lekens 1958:596b)

Sango:

- (4) *zò só mbi bá bíri la*
 person that 1S see yesterday TOP³⁷

Subordination and complementation. Ngbandi has several subordinators that have been replaced in Sango: e.g., *se ... ko* ‘if ... then’ by *tòngà nà*; and *me* ‘as’ by French *comme*, which, however, may have entered the language only recently (T286). The loss of the complementizer (or *déclaratif*) *yá* (whose lexical meaning is ‘thus’) is especially remarkable, given the fact that it is obligatory after verbs of saying and is used in other constructions where saying is implied.

In Sango, however, the verb *tènè* 'to say' is followed by the cited discourse. (Even though the verb has been reduced in Bangui to *tè*, I have not yet found evidence of the grammaticalization of this verb as a complementizer.)³⁸

Ngbandi:

- (5) *lò nè hé mbi yá*
 3s say give 1s say
 'he said to me (that) ...' (Lekens 1958:649a)

Sango:

- (6) *lò tènè nà mbi à-tènè ...*
 3s say PREP 1s SUBJ+say)
 'he said to me (that) ...'

Lexicon

The lexicon of Sango is very much reduced, as one would expect in a pidgin.³⁹ This fact is established in three ways.

(1) In using the 100-word list of allegedly change-resistant words established by Morris Swadesh for glottochronology, we found, as noted above, that there is a striking loss of Ngbandi words in Sango. And a reduction of 43.8 percent has been established in using the 500-word list prepared by Joseph E. Greenberg for Africa, when only 219 equivalents were obtained in 1969 in interviewing four persons in a Ngbaka village, one from each of four age groups (Roulon 1972:157).⁴⁰

(2) A more realistic but still somewhat arbitrary comparison of 150 monomorphemic Sango words under the letter 'B' with those in Ngbandi reveals the loss of about 94 percent of the vocabulary. The arbitrary restriction to only one part of the comparable lexicons was imposed by the size of the Ngbandi corpus (Lekens 1958, with 1,072 pages; Lekens 1955 was also used). Under 'B' 578 Ngbandi words were identified. To arrive at a similar comprehensive list of Sango was difficult. It required, first, the examination of eight dictionaries or word lists, dating from 1911 to 1995. This task required, in addition, having to make judgements about what could be accepted as a Sango word. (a) For example, the dictionaries of Bouquiaux et al. (1978) and SIL (1995) include words of Yakoma origin recently introduced, in my opinion, by an educated elite that seeks to enrich Sango's vocabulary.⁴¹ Three nouns of attested Ngbandi origin were deleted because of their source. For example, the word for 'victory' (*bèndà*) occurs only in a small dictionary published by the

Société Internationale de Linguistique (SIL) and prepared by Central Africans.⁴² (b) Moreover, many of these words occur only in a single source: e.g., 44 or 29.3 percent of the 150 words beginning with 'B' compiled by myself occur only in dictionaries prepared by members of Baptist Mid-Missions. Therefore, a more rigorous procedure than the one I have employed might result in a *greater* percentage of loss.

An analysis of the kinds of words found in these corpora is equally revealing of pidginization.⁴³ The *loss* of Ngbandi words in Sango, when classified by part of speech, reveals the following percentages: Connectives 100; Adverbs, 99; Adjectives, 98; Nouns, 93; Verbs, 92.

(3) In a corpus of tape-recorded extemporaneous speech of 37,217 words collected in 1962 (the only documentation of different words found in a large corpus until that which I began to collect in Bangui in 1988), there were only 489 irreducible lexical items that were not French in origin: in other words, excluding what might be considered compounds (Samarin 1966:188fn).⁴⁴ Even with such a corpus one is not certain that all words are widely used: some of my subjects might have used a word from their own languages for the nonce. (On this type of self-correction, see Samarin 1966.)

My view is that the total number of words in Sango is about one thousand, which was my estimate forty years ago (Samarin 1961) and that of Bouquiaux (1969:64, cited by Manessy 1977:145 and 1995:79 as "le vocabulaire usuel"). It has even been claimed that the average basic vocabulary necessary for someone to correctly use Sango is only a little more than two hundred words (Roulon 1972:157; cited by Manessy 1977:145).⁴⁵

In addition to this massive loss of vocabulary in Sango, many common Ngbandi function words were replaced or came to be used in a different manner. Of the many examples, here are only five.

(1) The use of the clitic *ngá* to signal a question requesting confirmation (in French, *est-ce que*) was lost in Sango, which uses an overall rise in pitch. Recently, however, people have begun using preposed *èsí* (from French *est-ce que*) for this purpose: Ngbandi *mò mú sà ngá* (2S receive meat INT); Sango (*èsí*) *mò wàrà nyàmà?* 'Did you get any meat?'

(2) The Ngbandi word *nà* 'who' (*ánà* in the plural) was replaced in Sango by *zò wà* (person who?) for 'who?'

(3) The simple Ngbandi *wà* 'where?' has been replaced by *nà ndò wà* (at place what?): Ngbandi *lò wà* (3S where?) (Lekens 1955); Sango, *lò kè nà ndò*

wà (3S be at place what?) ‘where is he/she?’⁴⁶

(4) For ‘how?’ *síá* (T511) was replaced by the phrase *nà lége nyè* (at way/road what?).

(5) The function of *nà* ‘with’ has changed so much that it affects the grammar of the language. While still being used with this meaning in Sango, it is also used as an all-purpose preposition. For example: (a) Benefactive after certain verbs: in Ngbandi the verb ‘give’ is not followed by a preposition, in Sango it is; (b) Locative after some verbs and in certain locutions: in Ngbandi ‘I went to the garden’ is expressed as *mbi gwè yaka* (1S go garden) (T502), but in Sango (as noted above) obligatorily as *mbi gwè à yaka*; (c) In Sango, but not in Ngbandi, *nà* is used to connect clauses.⁴⁷

In this section I have been concerned with demonstrating only that Ngbandi’s lexicon has been drastically reduced in Sango. Nonetheless, Sango’s vocabulary remains predominantly Ngbandi in origin.

4. Conclusion

My argument for recognizing that Sango is a pidgin will most certainly be accepted by all those who understand what is generally meant by the word *pidgin*. It falls into a type of language characterized by being (a) new and (b) remarkably reduced by comparison with its source language (and with any other natural language, for that matter).⁴⁸ This is not to say that it meets all the definitions of pidgins (for there are many definitions and many kinds of alleged pidgins), and it certainly fails to have what are claimed to be typical features of pidgins in general. Note the fifteen inventoried recently in Table 1.

Table 1. General features of pidgin grammars according to Sebba (1997:39-40), checked against those of Pidgin Sango. Key: + = true of Sango, 0 = not true of Sango, NR = not relevant because the lexifier language Ngbandi (and some if not all of the related and co-territorial languages) already have this feature.

1. No definite or indefinite article	0
2. No copula	0
3. Tense, aspect, modality and negation marked externally to the verb	+
4. No complex sentences	0
5. No passive forms	NR
6. Very few or no inflections	0
7. Analytic constructions used to mark possessive	NR

8. Avoidance of “highly marked” sounds	0
9. Simple syllable structure	NR
10. Tone is not used to distinguish words	0
11. Preference for semantic transparency	NR
12. Small vocabulary	+
13. Very small inventory of prepositions or postpositions	NR
14. Preference for short words	NR
15. Derivational morphology is not well developed	NR

Even though these features are listed as comprehensive and illustrative, and not criterial, they barely diagnose Sango as a pidgin. There are only two plusses, and one of them has to be qualified. Thus, although it is true of Sango, as it is alleged of pidgins in general (Sebba 1997:39), that “Tense, aspect, modality, and negation [are] marked externally to the verb” (that is to say, they are signalled lexically and not morphologically, and presumably in contrast with the base language), the statement applies only with respect to negation in Sango, but since it is also external to the verb in Ngbandi the (+) is a qualified one.⁴⁹ As has been already noted, the other grammatical categories are expressed tonally in Ngbandi.

Although I make my claim about Sango’s being a pidgin for reasons both linguistic and nonlinguistic, I would be comfortable in doing so for *only one reason* in spite of the claim that “it is very difficult [note the qualification] to find a single feature which occurs in *all* [italics in the original] pidgins without exception” (Sebba 1997:97). That “single feature”, I believe, is an extremely reduced lexicon. Is there, really, any natural language in the world with only about one thousand words? It is conceivable, of course, that there be languages with about that many roots, but they would be languages with the grammaticalized means for producing an infinite number of ‘words.’⁵⁰

I do not propose that a small lexicon is the *only* characteristic of a pidgin, only that it is the most salient and implicationally the most important. It would, however, be only the third of a set of the characteristics, beginning as follows:

1. A pidgin is a language.
2. It is a lingua franca.⁵¹
3. It has a very reduced lexicon.

Put in another way, I would suggest that if linguistic features were weighted in a diagnostic test, a reduced lexicon would have the greatest weight.

Notes

1. This paper was written in 1997. In 1887 Alphonse van Gele, in the name of Leopold II, King of the Belgians, arrived with a few whites and more numerous Africans at the headwaters of the Ubangi river, the birthplace of Sango (see map). The latter consisted of twelve Zanzibari, five Bangala, two Zulu, two ‘boys’ (that is, servants), and twenty-one ‘natives’, possibly so-called Bangala or conscripts from the Belgian pool of liberated captives (Samarin 1989a:147). For more information about van Gele see Cuypers (1960).
2. As some others before me, I will use Ngbandi as a cover-term for the language — which consists, we are informed, of a cluster of mutually intelligible dialects (Boyeldieu 1982) — except, when the need arises, for names of local varieties on the Ubangi and Mbomu rivers: e.g., Dendi, Yakoma, Sango.
3. This is not a rhetorically automatic qualification. The highest percentage of people who do not know Sango, or do not know it well, would in my opinion be found among Muslim immigrants from other African countries.
4. This conclusion, however, is based only on a sample of households. The percentages appear to be somewhat different for different parts of the city and for children with different ethnic histories. In other words, nativization in Sango is probably variable throughout the country.
5. Attributed to me is the claim that “Gbanziri was the *lingua franca* in the Ubangi bend before Sango emerged” (Pasch 1997:fn7). In the work cited (Samarin 1984/1985:341) all I had said was that in the small area where Gbanziri lived on the Ubangi river, to the east of Bangui, “their language *might* have served as the basis of some kind of *jargon*” and that it was “quite possible, and indeed quite likely, that Gbanziri was being *pidginized* by its foreign users, as was happening, we must assume, in the east in the dialects of [vernacular Sango]” (emphasis added here throughout). Indeed, one interpretation of a colonizer’s memoir would document this event (see Samarin 1982b:413).
6. Although Sango is listed in Holm (1988) as one of the pidgins and creoles of the world (see map on pp. xvi-xvii), it is cited in the index as “restructured Sango.” The index cites only one page for *restructuralization*, and for *restructuring* it advises the reader to see *creolization* and *pidginization*, which are not helpful for understanding this concept.
7. However, André Jacquot, one of the first French linguists to comment on Sango (Jacquot 1961), did refer to Sango as a pidgin, noting its “simplification” and “impoverishment” in personal communication (10 October 1959).
8. Manessy, however, would be willing to consider Sango a pidgin on sociolinguistic grounds (1977:132, cf. 130; 1995:57). This position is reinforced in the following, where, while acknowledging Sango’s use by people who (in my words) differed ethnolinguistically, Manessy & Wald (1984:69-70) say that Sango’s functions are not merely those of a contact language, because it has acquired prestige in being a language of wider communication and because of its use on the radio: “... les fonctions de cette langue ne se réduisent pas pour autant à celles d’un pidgin de contact: le prestige qu’acquiert cette langue dans l’ouverture d’un champ de communication plus large, son usage à la radio, sa place dans la vie publique depuis qu’elle est langue nationale (1964) et d’autres facteurs encore, rendent son acquisition pratiquement indispensable en milieu urbain et de plus en plus

- fréquente et désirable chez les ruraux.” Although these two scholars do admit that Sango issued from a contact language whose origins date “probablement” to the beginnings of colonization (*ibid.* 69fn), its linguistic nature is not specified. Manessy’s position is also ambiguous if not contradictory, because in the paper already cited he says that pidgins — not excluding Sango — will, among other languages, be used illustratively (Manessy 1977:132). Subsequently he recognizes certain pidginized features of Sango but explains their origin in language contact in an urban setting (1990:13, 15, 16, 18), apparently having been led to believe that the features he discusses arose first in Bangui. It is certainly true that a new variety of Sango is emerging in Bangui, the nation’s capital, but these changes have nothing to do with pidginization. Manessy finds that several changes that occur in urban varieties of vehicular languages are similar to those found in pidgins (1995:58, 59, 61, 63, 64, 65, 68, 75): e.g., analytical compounds, periphrasis, loss of inflection, simplification, “rendement maximal au moindre coût”, “principe d’économie”, loss of marked phonological features, phonological variation, open syllables, “univocité paradigmatique”, loss of grammatical markers, etc. (a view shared by Calvet 1992:14).
9. Cf. Pasch (1997:214, citing Diki-Kidiri 1982): “some variant of the Yakoma language is likely to have served as a vehicular language throughout the Ubangi bend ... they used Yakoma as a *lingua franca*”.
 10. Morrill, however, alleges that Diki-Kidiri’s view is just the contrary: that in the precolonial period “the Sango and the Yakoma tribes did *not* predominate on the full length of the Mbomu and Ubangi Rivers” (1997:157, emphasis added).
 11. The notebook in which A. van Gele recorded his purchases of ivory contains nothing about the Dendi. See the Alphonse Vangele archives in the library of the Musée Royal de l’Afrique Centrale, Tervuren, Belgium: B4, Booklet (8x14cm), inside cover ‘Ivoire 1888’ p. 1, “Carnet des achats d’ivoire. Rivière: Doua. 1er cahier.”
 12. Morrill claims to have “uncovered irrefutable evidence of Sango being spoken in Bangassou [on the Mbomu river] in April 1893” in words attributed to the sultan Bangassou by what the calendar would identify as a green-horn representative of the État Indépendent du Congo (1997:166; for lack of preparedness among these men see Samarin 1982b:414ff). By this time, I demur, given the aggressiveness of the Belgians in their pursuits, it would not be surprising at all that the emerging pidgin was spoken by some Africans at this site. After all, Alphonse van Gele, among others, had been active in this area. This ungrammatical concoction of words, however it was acquired and whatever may have been its communicative purpose, is surely no evidence of a stabilized language, as alleged (Morrill 1997:165). Moreover, this is not the first documentation of what might be the vehicular language: for others, and an earlier one, see Samarin (1982a).
 13. Nowhere in my reading did I ever find a single reference to an indigenous market in the Ubangi river basin. The French, and undoubtedly the Belgians, had to set them up in order that their men might buy provisions on a regular basis. Zigba (1995) arrived at the same conclusion.
 14. Paris, Archives Nationales Section Outre-Mer, Papiers Liotard 213 Mi [microfiche], 2, Journal 28 août 1891 au mai 1892 (“notes recopiées par la suite par un autre [...]”); 213 Mi 2(2), “Notes pour le rapport de fin de campagne (postérieur à 1894)”. For more on Liotard see de Mazières (1982).

15. Without historical documentation the Sango are said to have been the first to enter into contact with Europeans, serving as canoers (Bouquiaux et al. 1978:20-21), an assertion developed by Pasch (1997:215). In fact, the history of the colonization of the region demonstrates otherwise (Samarin 1989a).
16. My work, on the other hand, has led to explanations of the way two important words entered Sango: *kè* (its common pronunciation, but commonly written *yèkè*) 'to be' and *nginzà* 'money', the first apparently from Kikongo and the latter from some as-yet unidentified west African language (Samarin 1986a, 1989b). With respect to the first, there remains the possibility that either the copula is derived from Swahili or that the innovation involved the convergence of Swahili and Kikongo.
17. For the time being, consensus in believing in the existence of a kind of language identified as pidgin is being assumed. Morrill also recognizes the existence of pidgins (personal communication), but denies that Sango is one of them.
18. Thornell (1997:30) rejects my argument for the emergence of Sango in about two decades on the grounds that this "period of time [roughly between 1887 and 1911] seems too short." On the origins of pidgins, however, it has been demonstrated that some have arisen in a couple of decades. On Sranan, for example, see Voorhoeve (1964, 1971, 1973); on Chinook Jargon (a pidgin in spite of its name) see Samarin (1986b).
19. It seems necessary to be reminded of the distinction that must be drawn between pidginization as a *process* and a pidgin as a *product* (Samarin 1971).
20. I say "usually," because I can document the important role children played as language teachers and servants for Europeans — children who must have been involved, as adults were, in the creation of Sango.
21. The process described here is documented by one who was a member of a French expedition in Haut-Oubangui in 1892: "Nos hommes avaient peu à peu réuni quelques mots à l'aide desquels ils avaient constitué une sorte de patois, qui leur permettait jusqu'ici de se faire facilement comprendre par des indigènes" (Brunache 1894:206). This was probably Sango, for a lingua franca had already been recognized by whites. For a discussion of this text see Samarin (1985) and (1989c).
22. Long ago I characterized a pidgin as a language that "traces its lineage to at least one natural language" (1962:56).
23. Earlier (1986:299), in listing three kinds of lingua francas, Calvet includes "a language specifically created for communication purposes," citing Munukutuba, but then, "and of course, Pidgins", as if citing a different type of language but with no example. Calvet finds fault with my characterizing Sango as a creole in one of my early works (1992:10; see Samarin 1966); however, while my language could have been more explicit, my assessment of Sango's status could have been found in other publications cited at that time.
24. On studying these words again, this time with Lekens (1958), which I did not have on the first occasion, I find sixty-four unarguable retentions of Ngbandi words in Sango, and these include four that are somewhat problematic: whereas Swadesh gives a word in a nominal or adjectival form (e.g., *full*), the equivalents in Ngbandi and Sango are verbs.
25. The politics of dictionaries in new nation-states is not negligible. With respect to Sango this is seen in one politically oriented survey of dictionaries (Pénel 1979).

26. It should be noted that the following observations are based on the language that both Diki-Kidiri and I learned in the rural areas, for I am writing about Pidgin Sango. I am also writing about the *language*, not a special variety being deliberately created and advocated in recent years, a variety being 'engineered' by the Central African élite (Samarin 1998b).
27. It should be noted that Sango is a tone language, tone having been marked in my publications on Sango. It is surprising, therefore, to find in a very recent publication that one of the characteristics of pidgins in general is that they do not use tone to distinguish words (Sebba 1997:39). Moreover, "Those pidgins that are known to have lexifiers which are tone languages are *not* tone languages themselves. This is true of ... Pidgin Sango (whose lexifier...is tonal) ..." (Sebba 1997:48). And Sebba goes on to cite Bakker (1994:35) for having pointed out "that nearly all the speakers of Pidgin Sango speak tone languages natively [because, I might add, all languages of the Ubangian family are tonal as are the Bantu ones in the country] and Sango itself (the non-pidgin variety) is tonal" (the quotation marks enclose Sebba's statement, and italics were added above). This statement seems to imply that the pidgin variety of Sango is *not* tonal. Bakker has been misinformed. Every vowel in Sango carries either a low, mid, or high tone — contour tones being analyzed as occurring on sequences of vowels — and in contracted forms in Bangui today even some consonants carry tone. Moreover, there are a number of minimal pairs distinguished by tone alone: e.g., *kwá* 'death, corpse,' *kwa* 'hair,' *kwà* 'work [noun].' These facts were published long ago and more recently in Walker & Samarin (1997). Moreover, given the fact that some of the people involved in the 'creation' of Sango spoke languages that are *not* tonal (such as those who spoke Swahili [although some of these may have had native languages that were tonal] and Wolof) and that others spoke languages that used tone in a different way (such as Bambara, spoken also in Senegal) or had only two levels of tone instead of three, it is remarkable that Sango's lexical use of tone is so close to that of its source language.
28. Movement forward with the deletion of the preposition *tí* 'of', incidentally, is one of the variants of this preposition in Bangui today: e.g., *tère tí lò* (body of 3s) > *tère lò* 'his/her body.' The social distribution of this variant in Sango is under study.
29. Although this observation has not yet been documented in print, see Samarin (1991). One possible influence of Ngbandi may be in the absence of the preposition *nà* in the speech of young people in Bangui, especially, if not always, with certain verbs (the matter is being studied). Here, for example, is one from a popular song: *lò du [nà] n̄zàlà* (3S sit [with] hunger) 'he/she is hungry.' This deletion occurs in Ngbandi, for Lekens observed that this was a word that was not infrequently deleted, its tone, however, moving to the preceding lengthened vowel (1958:618a; for an example see Lekens 1955:123a). In Sango there is no lengthening of the preceding vowel.
30. By contrast, Diki-Kidiri (1977:727) characterizes Sango as a language with an economical morphology ("une langue à morphologie économique", which might better have been characterized as a "generously dis-endowed language"), typologically similar to Mande languages, like Bambara.
31. If tone is used grammatically, as Diki-Kidiri claims (1988), I believe that the speakers are of Ngbandi ethnicity and others who mimic them. I was participant in a conversation (in Sango) in which Diki-Kidiri told a Central African university student in France about this usage. With a surprised look she said, "I've never heard that." In my huge corpus of tape-

recorded extemporaneous speech collected in Bangui since 1988 a high tone on the third person singular pronoun or the verbal subject marker occurs rarely. A count of their occurrence in the transcriptions would confirm this impression. The grammatical function of tone is also asserted without exemplification by Roulon (1972:149). All examples in Sango are my own.

32. The missionaries were theologically motivated in wanting to make a distinction between *yé* 'to want, desire, agree to, consent, prefer' (their meanings in Ngbandi as well) and 'to love', as in "Love your neighbor as yourself."
33. I confess, however, that this is a conclusion based on observation, not on elicitation. I have not asked people who were not native speakers of Ngbandi to explain the difference in meaning between these two words, and I believe that the entry in Bouquiaux's dictionary for the derived form is another instance of its bias toward a conscious reconstitution of Sango. In any case, since this appears to be the only iterative verb in the dictionary, it can very well be considered nothing more than a vestigial item.
34. These are similar to what Mühlhäusler calls "phrase level lexical items" (1995:112), illustrated in Tok Pisin by *man bilong kros* (< *man belong cross*) 'a bad-tempered person', typologically similar to Sango's *zò tí ngònzò* (person of anger). Diki-Kidiri is therefore mistaken about the alleged productivity of the N + N construction with the "direct complement". From a corpus of 1,700 *tí*-phrases in extemporaneous speech from a good sample of Central Africans, I found twenty instances of the whole phrase *yángá (tí) dà* 'mouth of house'. Although nineteen occurred without the preposition, it should be noted that the speakers were predominantly of Ngbandi ethnicity, others of Ngbaka ethnicity. In this latter population I have come to expect Ngbandi features. In any case, a thirteen-year-old girl of Ngbaka ethnicity used the phrase twice: once *without* the preposition and once with a reduced form of it: namely, the consonant *n* with high tone. (Others use the same form, and I have also found *m* used in the same way for the preposition.) The analysis of the variable {*tí*} in tape-recordings of persons of Ngbandi and of Gbaya ethnicity is now under way.
35. Although I once attributed this negative to a Bantu language, not knowing of the last-mentioned word in Ngbandi, I may have been mistaken. In any case, negation has been greatly simplified.
36. The study of *só* by comparison with its use in Ngbandi is under way. These glosses are based on translations of Ngbandi into French by an educated native-speaker of Yakoma and on translations into English of Lekens' Flemish translations by native-speakers of Flemish.
37. What I have identified as a topicalizer is glossed by Lekens (in translation from Flemish) as 'there is, there are.' Its function in Sango, especially the Sango of Bangui, awaits study.
38. The parenthesized 'that' simply indicates its presence or absence in two kinds of discourse: indirect vs. direct. In the latter, both in Ngbandi and Sango a different pronoun is used. This is also true in Gbeya.
39. We expect a reduced lexicon because of what has been learned about all such new languages. It would serve no purpose to trace the history of statements with respect to pidgin vocabularies, but it should at least be noted that Reinecke discussed this topic sixty

- years ago (1937:139), where, in his typical and admirably cautious manner, he estimated “the number of [Chinook] Jargon words probably in general use at any one time and place at about 500.”
40. It is not clear how she interviewed the subjects, because she says “nous [meaning herself] n’avions pas de langue de communication” (Roulon 1972:139). Inquiries about a published form of Greenberg’s list proved futile.
 41. Bouquiaux’s includes many words from the ethnic language (identified as *sango riverain*) with no declared justification. (One of the collaborators is of that ethnicity.) Second, it includes many neologisms, some of which are identified by an asterisk, some not. (The other collaborator has been active in enriching Sango with neologisms. See, for example, Diki-Kidiri [1982]. This latter work is a bilingual text, with French on the even-numbered pages and Sango on the odd-numbered ones. It also includes Sango-French and French-Sango glossaries. For example, borrowing from Ngbandi, the author creates *piàlo* for ‘projet’ out of *pià* ‘égale, juste’ (ADJ/ADV) + *lo* ‘parole,’ neither of which occurs in Sango.) Third, it is arguable whether some of these are really words or phrases, and whether they are possibly idiosyncratic (see Samarin 1980 for a review).
 42. Like so many of these neologisms, this one is curious. In Ngbandi the word means ‘monstre’ and ‘énormité’ (Lekens 1955:188a; 1958:44b).
 43. In this *b*-corpus of Ngbandi the number of adjectives, adverbs, and ideophones is 68, 11 percent of the total. In Sango the number is only six, 4 percent. Several years ago a specific study was undertaken with respect to the use of ideophones in Sango by tape-recording extemporaneous descriptions of slides depicting the making of a clay pot with the coil method. Some of the subjects were Gbeya, who in Gbeya used, as was expected, a large number of them; those restricted to Sango did not have the vocabulary to be as precise. See Samarin (1979).
 44. This same corpus was used in my research project for Taber’s dictionary (1965), with a number of words added from other sources in consultation with my Central African assistant Simon Nam-bo-zui-na. Another corpus of tape-recorded speech of 51,781 “running words” (apparently meaning word tokens) was recently obtained in the western part of the CAR (Thornell 1997:53), but we have not been informed of the number of word *types* in it.
 45. In French: “On peut considérer que ce chiffre représente le vocabulaire de base moyen nécessaire pour qu’un individu manie correctement cette langue.” But this opinion is based, as has already been noted, on Greenberg’s 500-word list. Sango certainly has words that are not included in that list: e.g. *mè* (< Ngbandi *mé* ‘to beat’) ‘to mix flour and water into a dough’. It is surprising, therefore, to find it averred that the vocabulary of Sango and those of surrounding Ubangian languages “are more similar in size”, apparently meaning that although Sango’s vocabulary is small by comparison with European vernaculars, it is not much different from that of other Ubangian languages (Thornell 1997:94). By attributing 6,000 entries to the Sango dictionary by Bouquiaux et al. (1978), she implies that Gbaya and Banda have about that many words. Whatever the size of Sango’s lexicon, it must certainly be less than those of other Ubangian languages (I have a collection of almost 6,000 ideophonic adverbs alone for Gbeya). Moreover, she would have found a rich vocabulary in the dictionary of a dialect of Gbaya similar to if not mutually intelligible with the one at Carnot where she did most of her work (Blanchard &

- Noss 1982; Noss 1981); the later and much more comprehensive dictionary of Ngbandi than the one cited (Lekens 1958) would also have been helpful in her analysis. (The small number of words in my corpus led Einar Haugen to wonder if Sango was a language at all! See Samarin 1966.)
46. In Sango *wà* is an interrogative, additional to and sometimes interchangeable with *nyè*: e.g., *zò wà* and *zò nyè* (person what?) 'who?'
 47. I suspect the influence of one or more Bantu languages — possibly Swahili and what came to be known as Kituba — in the expansion of the use of this preposition, *at the time of the origin of Sango*, but this topic is beyond the scope of the present paper.
 48. The only attempt at devising a diagnostic tool for identifying a pidgin, based on a hypothesis suggested by Martin Joos, was the measure of hapax legomena in a language (Samarin 1971:120, followed by comment by H. A. Gleason, Jr.). I am not aware of its ever being criticized or approved. In any experiment, great rigor would have to be exercised in the selection of texts and great sophistication in the use of statistics. The paper just cited, incidentally, is where Gleason is also given the credit for having proposed the idea that "Pidginization should be seen as *any consistent reduction of the functions of language both in its grammar and its use*" (Samarin 1971:126, the words being mine and the italics in the original).
 49. Moreover, characteristics attributed to pidgins are always (I dare to make that assertion) hedged with qualifications. For example, in a recent survey of pidginization and pidgins the following words and phrases occurred: *as a general rule, characteristically, fairly typical, generally, have a tendency, in all likelihood, in most cases, is likely to, may, might, often, rather, relatively, seem to, tend to, typically, usually*. They were needed, I will admit, in the kind of discourse that was being produced. But for how much longer, and for what reasons, will they be necessary? Until, I suspect, there are adequate descriptions of pidgins and until linguists familiarize themselves with more than one or two of them.
 50. Although Mühlhäusler believes that circumlocution had a "vigorous presence in the early years" of Tok Pisin, he does not believe that they should be regarded as "full-fledged lexical items": e.g., *marasin bilong stopim bel bilong karim pikinini* (medicine of stop belly of carry child / medicine which prevents conception) for 'contraceptive pill' (Mühlhäusler 1985:68, 227, 228). In this same place, incidentally, he traces the history of the legendary circumlocution for piano.
 51. It must be clearly understood that phantoms identified as *PCs* are outside the purview of this argument. We are only discussing pidgins that are lingua francas. For the reader who may not be acquainted with the abbreviation it must be explained that it stands for "pidgins-and-creoles", as if spoken in one breath, as *ham-and-eggs* and *Fourth-of-July* in the United States. Of course, if the linguist is speaking seriously about pidgins, he or she might use the full phrase and probably say it slowly. In published material, on the other hand, one can find statements about *PCs* that really do not apply to pidgins in general. So common is the use of this abbreviation, that it has spawned *PC-ization* and other such neologisms in both English and French even though this purported portmanteau phenomenon has yet to be documented and argued. On lingua francas see Samarin (1987); see also Kahane & Kahane (1976).

References

- Bakker, P. 1994. "Pidgins". In *Pidgins and Creoles: An Introduction*, J. Arends, P. Muysken, and N. Smith (eds), 25-40. Amsterdam: John Benjamins.
- Blanchard, Y. (O.M.I.) and Noss, P.A. 1982. *Dictionnaire gbaya-français: Dialecte yaayuwée*. Meiganga, Cameroun: Centre de Traduction Gbaya.
- Bouquiaux, L. 1969. "La créolisation du français par le sango véhiculaire, Phénomène Réciproque". In *Le français en France et hors de France. 1: Créoles et contacts africains*, 57-70. Paris: Les Belles Lettres.
- Bouquiaux, L., Koboza, J.-M., and Diki-Kidiri, M. 1978. *Dictionnaire sango-français, lexique français-sango*. (Langues et Civilisations à Tradition Orale, 29.) Paris: SELAF.
- Boyeldieu, P. 1982. "Structures sociales et particularismes linguistiques en pays de langue 'ngbandi': Éléments pour une étude". In Boyeldieu and Diki-Kidiri, 13-80.
- Boyeldieu, P. and Diki-Kidiri, M. (eds) 1982. *Le Domaine ngbandi*. Paris: SELAF.
- Bruel, G. 1918. *L'Afrique Équatoriale Française: Le pays, les habitants, la colonisation, les pouvoirs publics*. Paris: Larose.
- Brunache, P. 1894. *Le centre de l'Afrique: Autour du Tchad*. Paris: Ancienne Librairie Germer Baillière & Cie.
- Calloc'h, J. (Révérend Père, C.S.E.). 1911. *Vocabulaire français-sango et sango-français: Langue commerciale de l'Oubangui-Chari, précédé d'un abrégé grammatical*. Paris: Librairie Paul Geuthner.
- Calvet, L.-J. 1974. *Linguistique et colonialisme: Petit traité de glottophagie*. Paris: Petite Bibliothèque Payot.
- Calvet, L.-J. 1986. "Trade function and lingua francas". In *The Fergusonian Impact: In Honor of Charles A. Ferguson on the Occasion of his 65th Birthday. Volume 2: Sociolinguistics and the Sociology of Language*, Fishman, J.A. Tabouret-Keller, A., Clyne, M., Krishnamurti, B. and Abdulaziz, M. (eds), 295-302. Berlin: Mouton de Gruyter.
- Calvet, L.-J. 1992. "Ce que la linguistique doit aux études créoles". *Études Créoles* 15(2):9-44.
- Crampel. 1891. Lettre de M. Crampel à M. Musy père. *Revue de Géographie* 29:141-143.
- Cuyppers, J.-P. 1960. *Alphonse Vangele (1848-1939)*. Bruxelles: Académie Royale des Sciences d'Outre-Mer.
- De Mazières, A.-C. 1982. *La marche au Nil de Victor Liotard: Histoire de l'implantation française dans le Haut-Oubangui 1891-1899*. Aix-en-Provence: Université de Provence.
- Diki-Kidiri, M. 1977. "Développement du sango pour l'expression du monde moderne: obstacles et possibilités". In *Les relations entre les langues négro-africaines et la langue française*. Paris: Conseil International de la Langue Française (CILF). Pp. 717-728.
- Diki-Kidiri, M. 1979a. "L'émergence du sango comme langue nationale centrafricaine". *Recherche, pédagogie et culture*. Vol. 8, N° 43, *Langue, Individu, Société*, pp. 36-39.
- Diki-Kidiri, M. 1979b. "Créations lexicales spontanées en sango à partir d'emprunts au français". In *Contacts de langues et contacts de cultures* 3, Caprile, J.-P. (ed.), 33-43. Paris: SELAF-ACCT.

- Diki-Kidiri, M. 1981. "L'expansion du sango en Centrafrique". In *L'expansion des langues africaines: Peul, sango, kikongo, ciluba, swahili*, Caprille, J.-P. (ed.), 29-42. LACITO-Documents, Afrique 8, série 'Contacts de Langues et Contacts de Cultures'. Paris: SELAF.
- Diki-Kidiri, M. 1982. "Variations libres entre phonèmes en sango, langue nationale de R.C.A. Étude et mesures". In Boyeldieu and Diki-Kidiri, 81-103.
- Diki-Kidiri, M. 1985a. "Le sango entre l'oralité et l'écriture". In *Passage à la lecture en Afrique*. AELIA (Association d'études linguistiques interculturelles africaines). Bulletin, Nu. 8, Décembre. Actes du colloque international sur le passage à la lecture en Afrique noire, 39-46. A.E.L.I.A., A. P.E.L.A., O.R.S.T.O.M., and Université de Bénin.
- Diki-Kidiri, M. 1985b. "Élaboration de terminologie scientifique dans une langue de tradition orale: le sango, langue nationale centrafricaine". *Discoss* 1:77-82.
- Diki-Kidiri, M. 1985c. "La création lexicale en sango". MS.
- Diki-Kidiri, M. 1986. "Le sango dans la formation de la nation centrafricaine". *Politique Africaine* 23:83-99.
- Diki-Kidiri, M. 1987a. "La situation du sango en République Centrafricaine". *Etudes de Linguistique Appliquée: Politiques linguistiques (études de cas)*, Boyer, H. and Dumont, P. (eds), 102-09.
- Diki-Kidiri, M. 1987b. "Le complétif direct et indirect en sango". In *La maison du chef et la tête du cabri: Des degrés de la détermination nominale dans les langues d'Afrique Centrale*, Boyeldieu, P. (ed.), 113-25. Paris: LACITO, Département Langues et Parole en Afrique Centrale (LAPAC)/Librairie Orientaliste Paul Geuthner.
- Diki-Kidiri, M. 1988. "Aspects, modes et temps en sango". In *Temps et aspects*, Tersis, N. and Kihm, A. (eds), 117-24. Paris: Peeters & SELAF.
- Diki-Kidiri, M. 1992. "Le sango en Centrafrique aujourd'hui". *Université* (May): 23-4.
- Gilman, C. 1979. "Lexical loss in Cameroon Pidgin". *Anthropos* 74:174-180.
- Holm, J. 1988. *Pidgins and Creoles*, Vol. 1. Cambridge: Cambridge University Press.
- Huttar, G. 1996. "'Basic vocabulary': light from creoles". Paper presented at the annual meeting of the Society for Pidgin and Creole Linguistics.
- Hymes, D. 1971. (ed.) *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press.
- Jacquot, A. 1961. "Notes sur la situation du sango à Bangui: Résultats d'un sondage". *Africa* 31:158-166
- Kahane, H. and Kahane, R. 1976. "Lingua franca: the story of a term". *Romance Philology* 30:25- 41.
- Kamanda, K. 1989. "La conjugaison en ngbandi (langue non bantou)". *Annales Aequatoria* 10:181-199.
- Lekens, B. 1955. *Dictionnaire ngbandi*. Anvers: Éditions de Sikkell.
- Lekens, B. 1958. *Ngbandi-Idioticon, II: Ngbandi en Frans-Nederlands*. Tervuren: Musée royal du Congo belge.
- Manessy, G. (trans. by Albert Valdman). 1977. "Processes of pidginization in African languages". In *Pidgin and Creole Linguistics*, Valdman, A. (ed.), 129-49. Bloomington: Indiana University Press.
- Manessy, G. 1990. "Modes de structuration des parlers urbains". *Actes du Colloque International 'Des Langues et des Villes', Dakar, 15-17 December 1990*. Paris: Didier Érudition.

- Manessy, G. 1995. *Créoles, pidgins, variétés véhiculaires: Procès et genèse*. Paris: CNRS Éditions.
- Manessy, G. and Wald, P. 1984. *Le Français en Afrique Noire: Tel qu'on le parle, tel qu'on le dit*. Paris: Éditions L'Harmattan.
- Morrill, C.H. 1997. *Language, Culture, and Society in the Central African Republic: The Emergence and Development of Sango*. Ph.D. dissertation, Bloomington, Indiana University.
- Mühlhäusler, P. 1985. "Patterns of contact, mixture, creation and nativization: their contribution to a general theory of language". In *Developmental Mechanisms of Language*, Bailey, C.-J. N. and Harris, R. (eds), 51-87. Oxford: Pergamon Press.
- Mühlhäusler, P. 1995. "Morphological expansion in pidgin languages". In *The Development of Morphological Systematicity: A Cross-linguistic Perspective*, Pishwa, H. and Maroldt, K. (eds), 107-20. Tübingen: Gunter Narr.
- Musy, M. 1890-1891. Correspondance de M. Maurice Musy, chef de poste à Bangui (Congo français), massacré par les indigènes. *Revue de Géographie* 27:375-81, 455-57 (1890); 18:64-68, 130-33, 211-14, 291-94; 29:62-66 (1891).
- Noss, P.A. 1981. *Grammaire gbaya, dialecte yaayuwée*. Meiganga, Cameroun: Centre de Traduction Gbaya.
- Pasch, H. 1997. "Sango". In *Contact Languages: A Wider Perspective*, Thomason, S.G. (ed.), 209-70. Amsterdam: John Benjamins.
- Pénel, J.-D. 1979. "Idéologies et dictionnaires sango". *Sendayanga ti Laso/Linguistique Actuelle*. Bulletin de l'Institut de linguistique appliquée de Bangui. N° 1, État des recherches linguistiques en RCA, 32-41.
- Reinecke, J. 1937. *Marginal Languages: A Sociological Survey of the Creole Languages and Trade Jargons*. Yale University doctoral dissertation. (Ann Arbor, Michigan: University Microfilm International.)
- Roulon, P. 1972. "Étude du français et du sango parlés par les Ngbaka-Ma'Bo (République Centrafricaine)". *Éthnies* 2:133-65.
- Samarin, W. J. 1953. *Learning Sango: A Pedagogical Grammar*. Bozoum, C.A.R.: Mission Évangélique de l'Oubangui-Chari. Mimeographed.
- Samarin, W. J. 1955. "Sango, an African lingua franca". *Word* 11:254-67.
- Samarin, W. J. 1958. "The phonology of Pidgin Sango". *Word* 14:62-70.
- Samarin, W. J. 1961. "The vocabulary of Sango". *Word* 17:16-22.
- Samarin, W. J. 1962. "Lingua francas, with special reference to Africa". In *Study of the Role of Second Languages in Asia, Africa, and Latin America*, Rice, F.A. (ed.), 54-64. Washington, D.C.: Center for Applied Linguistics.
- Samarin, W. J. 1966. "Self-annulling prestige factors among speakers of a creole language". In *Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference, 1964*, Bright, W. (ed.), 188-213. The Hague: Mouton.
- Samarin, W. J. 1967a. *A Grammar of Sango*. The Hague: Mouton.
- Samarin, W. J. 1967b. *Lessons in Sango. Volume 1: Basic Course in Sango*. Hartford, CT: Hartford Seminary Foundation.
- Samarin, W. J. 1970. *Sango, langue de l'Afrique Centrale*. Leiden: E. J. Brill.
- Samarin, W. J. 1971. "Salient and substantive pidginization". In Hymes, 117-40.
- Samarin, W. J. 1979. "Simplification, pidginization, and language change". In *Readings in Creole Studies*, Hancock, I.F. (ed.), 55-68. Ghent, Belgium: E. Story-Scientia.

- Samarin, W. J. 1980. Review of Bouquiaux, et al. 1978. *Journal of African Languages and Linguistics* 2:159-64.
- Samarin, W. J. 1982a. "Colonization and pidginization on the Ubangi river". *Journal of African Languages and Linguistics* 4:1-42.
- Samarin, W. J. 1982b. "Goals, roles, and language skills in colonizing central equatorial Africa". *Anthropological Linguistics* 224:410-22.
- Samarin, W. J. 1984a. "The linguistic world of field colonialism". *Language in Society* 13:435-53.
- Samarin, W. J. 1984b. "Bondjo ethnicity and colonial imagination". *Canadian Journal of African Studies* 18:345-65.
- Samarin, W. J. 1984/1985. "Communication by Ubangian water and word". *Sprache und Geschichte in Afrika* 6:309-73. See Samarin 1984c.
- Samarin, W. J. 1985. "Plurigenesis in pidginization". *Journal of Historical Linguistics and Philology* 2(1/2):100-19.
- Samarin, W. J. 1986a. "The source of Sango's 'be'". *Journal of Pidgin and Creole Languages* 1:205-23.
- Samarin, W. J. 1986b. "Chinook Jargon and pidgin historiography". *Revue Canadienne d'Anthropologie* 5:23-34.
- Samarin, W. J. 1986c. "Protestant Missions and the History of Lingala". *Journal of Religion in Africa* 16:139-63.
- Samarin, W. J. 1987. "Lingua franca". In *Sociolinguistics/Soziolinguistik* (Vol. 1), Ammon, U., Dittmar, N. and Mattheier, K.J. (eds), 371-4. Berlin: Walter de Gruyter.
- Samarin, W. J. 1988. "Le comportement linguistique des colonisateurs dans le Haut Congo français". In *Dialoguer avec le léopard? Pratiques, savoirs et actes du peuple face au politique en afrique noire contemporaine*, Jewsiewicki, B. and Moniot, H. (eds), 49-73. Paris: L'Harmattan.
- Samarin, W. J. 1989a. *The Black Man's Burden: African Colonial Labor on the Congo and Ubangi Rivers, 1880-1900*. Boulder, Colorado: Westview Press.
- Samarin, W. J. 1989b. "The colonial heritage of the Central African Republic: a linguistic perspective". *The International Journal of African Historical Studies* 22:697-711.
- Samarin, W. J. 1989c. "Language in the colonization of central Africa". *Canadian Journal of African Studies/Revue canadienne des études africaines* 23:232-48.
- Samarin, W. J. 1990. "The origins of Kituba and Lingala". *Journal of African Languages and Linguistics* 12:47-77.
- Samarin, W. J. 1991. "An incipient ethnic model for urban Sango". *Anthropological Linguistics* 33:166-82.
- Samarin, W. J. 1993. "Suffixation from phonological change in creolized Sango". Paper presented at the annual meeting of the Society for Pidgin and Creole Languages.
- Samarin, W. J. 1995. "The urbanization and creolization of Sango". Paper presented at the Twenty-sixth Annual Meeting on African Languages, University of California at Los Angeles.
- Samarin, W. J. 1997. "The creolization of pidgin morphophonology". In *The Structure and Status of Pidgins and Creoles*, Spears, A.K. and Winford, D. (eds), 175-216. Amsterdam: John Benjamins.
- Samarin, W. J. 1998a. "The creation and critique of a Central African myth". *Revue Française d'Histoire d'Outre-mer*, tome 85, no° 318, 55-81.

- Samarin, W. J. 1998b. "External forces in remaking pidgin Sango". Paper presented at the conference 'La Coexistence des Langues dans l'Espace Francophone: Approche Macrosociolinguistique' in Rabat, Morocco.
- Sebba, M. 1997. *Contact Languages: Pidgins and Creoles*. New York: St. Martin's Press.
- Société internationale linguistique (alias Summer Institute of Linguistics) and Institut de Linguistique Appliqué. 1995. *Petit dictionnaire sango*. Bangui, Central African Republic.
- Taber, C.R. 1965. *A Dictionary of Sango*. (W. J. Samarin, project director). Hartford, Connecticut: Hartford Seminary Foundation.
- Thomason, S.G. 1981. Review of *The Genesis of Language*, ed. by Kenneth C. Hill. In *Studies in Second Language Acquisition* 3:237-49.
- Thornell, C. 1997. *The Sango Language and its Lexicon*. Lund: Lund University Press.
- Toronzoni N.-N. T. N. 1989. *Description du ngbandi, langue oubanguienne du nord-ouest du Zaïre*. Dissertation présentée en vue de l'obtention du grade de docteur en linguistique afraine, Bruxelles: Université libre de Bruxelles.
- Voorhoeve, J. 1964. "Creole languages and creolization". *Symposium on Multilingualism (Brazzaville, 1962)*, 233-42. London: Commission de Coopération Technique en Afrique.
- Voorhoeve, J. 1971. "Varieties of creole in Suriname: Church Creole and pagan cult languages". In Hymes, 305-15.
- Voorhoeve, J. 1973. "Historical and linguistic evidence in favor of relexification theory in the formation of Creoles". *Language in Society* 2:133-45.
- Walker, J. and Samarin, W.J. 1997. "Sango phonology". In *Phonologies of Asia and Africa (including the Caucasus)*, Vol. 2, Kaye, A.S. (ed). Winona Lake, IN: Eisenbrauns. Pp. 861-80.
- Zigba, D.B. 1995. *Navigation et échanges commerciaux chez les riverains du bassin supérieur de l'Oubangui (XIXe et XXe siècles)*. Thesis for docteur de l'université de Paris X, Nanterre.

Optimality Theory, the Minimal-Word Constraint, and the Historical Sequencing of Substrate Influence in Pidgin/Creole Genesis

John Victor Singler¹

1. Introduction

In this article I argue that the perspective of a constraint-based theory of phonology offers insights into the nature of substratal and universal influences upon pidgin/creole (PC) phonology. I then argue that the use of such a theory enables one to see that the chronological sequence of input from different substratal sources into a PC plays a vital role in determining the ultimate character of the PC's phonology. It isn't enough to know what the substratal languages are; one must also know when, relative to one another, these languages would have been present during the course of PC genesis. (Singler 1986 et seq., Arends 1986 et seq., and Chaudenson 1979, 1992 all make the point that PC genesis is a gradual process.)

2. Vernacular Liberian English

Vernacular Liberian English (VLE) spans a continuum (cf. DeCamp 1971, Singler 1997); its basilectal endpoint is highly pidginized. VLE is distinct from the other Liberian Englishes, i.e. Kru Pidgin English, Liberian Settler English, and Standard Liberian English. Of these other Englishes, Kru Pidgin English is the variety spoken in the small home villages of the "Kru sailors" along the southeastern coast (cf. Singler 1988, 1990), and Liberian Settler English is the vernacular of the descendants of African-Americans who immi-

grated to Liberia in the nineteenth century (cf. Singler 1989, forthcoming). Liberian Settler English has exerted extensive influence on VLE, particularly in the latter's mesolectal and acrolectal range; nonetheless, VLE is most accurately characterized as a modern descendant of the English-lexifier pidgin that was already spoken along this portion of the West African coast when the Settlers arrived (Singler 1997).

In the present study I distinguish between Coastal and Interior varieties of VLE. My discussion begins with an examination of Interior VLE. At first I will consider verbs in the Interior variety and then examine other word classes there as well. As will be seen, at first glance the coda consonants of verbs in Interior VLE appear to require an analysis different from the coda consonants of other word classes; however, I argue for a unitary analysis for all coda consonants in Interior VLE. Subsequently, I relate the treatment of coda consonants in Interior VLE to their treatment in Coastal VLE.

A crucial difference between Coastal and Interior VLE is that, no matter how basilectal the grammar of Coastal VLE, epenthetic vowels are virtually never added word-finally. Thus, the passage in (1) from a coastal speaker in the basilectal range contains no epenthetic vowels. In contrast, in Interior VLE, epenthetic vowels at the end of verbs occur frequently though not invariably, as illustrated in the passage in (2). In (1) and (2) the top line represents the speaker's actual pronunciation, while the second line displays the English word that is the source of the VLE word.²

(1) Coastal VLE

Wi ga gene pi, grāhɔ, bʌɛ, ẽ ɔ kẽ i rai. So
 we got guinea pig groundhog bird and all can eat rice so
enitam wi me awa fam, dẽ eritẽ es chakla.
 anytime we make our farm then everything is chakla
 'We have guinea pigs, ground hogs, birds, everything; and they all
 eat our rice. Anytime we make a rice farm, everything gets messed
 up.' (Bassa Porter)

(2) Interior VLE

Bikɔ awɔke, no mone. Das tẽ a livi tu nafisetese,
 because I work no money that's thing I leave to 1966
a livi di wɔk, tu kɔ tu ma hõ.
 I leave the work to come to my home
 'Because I worked but couldn't save any money. That's why I left
 in 1966, I left the work to come back home.' (Lofa Overseer)

(The paragogic vowel in VLE is ordinarily [i] or [e], depending on the height of the vowel preceding it.)

As in Singler (1991, 1996a), the data for the present study of Interior VLE come from sociolinguistic interviews with elderly men in Borkeza, Lofa County, Liberia. All of them are native speakers of Loma.³

3. Phonological Theory and the Role of the Substrate

While it is widely acknowledged that substratal languages play a pivotal role in shaping PC phonology, the specifics of that role are often left vague. This imprecision or lack of understanding is in part a consequence of the theoretical model that has been employed, specifically the SPE-type approach to phonology with its crucial reliance upon language-particular rules.

Data from Singler (1991), a study of the word-final consonants of verbs in Interior VLE, show ways in which the presence of substratal influence upon a PC is obscured rather than captured by an SPE model (i.e. based upon the 1968 classic *The Sound Patterns of English* by Noam Chomsky and Morris Halle). The primary substrate languages for the VLE of the interior are all languages of the Mande branch of Niger-Congo: the Southwestern Mande languages Loma, Kpelle, and Bandi; and the Southeastern Mande languages Mano and Dan. (None of the five are mutually intelligible.) In all five of these Mande languages, every syllable must be vowel-final. Most syllables in Interior VLE are vowel-final, too, even if the corresponding English syllable contains a coda consonant.

Making a syllable vowel-final in Interior VLE when the corresponding English syllable is not vowel-final is achieved by one of the set of phonological processes illustrated in (3), specifically deletion, resyllabification, or paragoge:

- (3) Processes by which syllables are made vowel-final in Interior VLE:
- | | | | | |
|----|-------------------|-----------|-----------------|----------------|
| a. | deletion | /ripit/ | <i>ri.pi</i> | ‘to repeat’ |
| b. | resyllabification | /tek ewe/ | <i>te.ke.we</i> | ‘to take away’ |
| c. | paragoge | /tek/ | <i>te.ke</i> | ‘to take’ |

It is through these processes that Interior VLE brings English lexical items into conformity with Mande phonotactics, and rules particular to Interior VLE can be devised in an SPE model to express each one of them. While it certainly seems reasonable to think of this as substratal influence at work, in fact

deletion, resyllabification, and paragoge themselves are not substratal phenomena. Loma, for example, has no paragoge, no resyllabification, and only rarely has any deletion. In other words, if we approach the study of PC phonology from the perspective of rules and derivations, we have to ask why a PC seems to be inventing the apparatuses by which phonological transfer from the substrate takes place. If this is substratal influence, why is it that carryover from a speaker's first language into a PC has to be accomplished by improvised processes? Further, all three of the processes in (3) succeed in eliminating coda consonants: on what basis is a particular process selected in a given instance? Resyllabification is limited to cases where there is an available slot in the following onset, but when does deletion occur and when paragoge?

In the 1990's, several constraint-based theories of phonology emerged, including Optimality Theory (OT), developed in Prince & Smolensky (1993) and McCarthy & Prince (1993a, 1993b, 1994). In what follows, I make use of OT to reformulate the role of the substrate in shaping PC phonology. Whereas a rule-based model seems to require the improvisation of the mechanism by which substratal influence enters the PC, in an OT formulation, critical elements of the phonology of the substrate go directly into the PC. To illustrate this, I begin by presenting a brief summary of the basic principles of OT, and then identify a crucial site for substratal influence upon a PC in OT's conceptualization of phonology.

Essentially, OT is "a nonderivational, surface-oriented theory" (Moon 1996:3). In place of rules and derivations, it considers all potential surface outputs, thereby providing a metric for the selection of the optimal output. Among the major principles of OT are these (drawn verbatim from Archangeli 1997:16 and ultimately from McCarthy & Prince 1993a, 1993b, 1994):

1. *Universal Grammar includes*
 - a. a linguistic alphabet
 - b. a set of constraints, **CON**
 - c. two functions, **GEN**(erate) and **EVAL**(uate)
2. *The grammar of a particular language includes*
 - a. basic forms for morphemes (from which inputs are constructed)
 - b. a **ranking** for the constraints in **CON**
3. *For each input,*
 - a. **GEN** creates a candidate set of potential outputs
 - b. **EVAL** selects the optimal candidate from that set (Archangeli 1997:16)

In this model, all constraints are seen as being part of Universal Grammar; at the same time, a given constraint is violable. It is the ranking of constraints that is language-particular. Put another way, a crucial way in which one language differs from another is that each has its own determination as to which constraints' violation it is most important to avoid; it is the language-particular ranking of constraints that expresses this.

An examination of the influence of Liberian Mande languages upon Interior VLE demonstrates that the primary locus of substratal influence lies in surface structure conditions, i.e. in constraints on the output. Thus, OT makes it possible to characterize PC phonology in a straightforward way: the PC's input comes (in the usual case) from the lexifier language, the constraints that govern the PC's output are part of Universal Grammar, and the ranking of those constraints in the PC comes in part or entirely from the PC's substrate languages.

4. Word-final Consonants in Interior VLE

I will illustrate this by presenting an OT analysis of word-final consonants in Interior VLE. In keeping with current practice within OT, I assume the Correspondence Theory of McCarthy and Prince (McCarthy & Prince 1995, McCarthy 1995) and the Faithfulness Constraints that follow from it. These constraints are given in (4):

(4) Faithfulness Constraints:

MAX Every segment of the input has a correspondent in the output (i.e. no phonological deletion).

DEP Every segment of the output has a correspondent in the input (i.e. no epenthetic segments).

Lombardi (1995, 1998) argues that Faithfulness Constraints exist for features as well (and not simply for segments). In the spirit of Lombardi, I follow Weinberg (1996) in distinguishing between DEP for consonants and DEP for vowels; I set this revision out in (4'):

(4') DEP-C Every consonant of the output has a correspondent in the input (i.e. no epenthetic consonants).

DEP-V Every vowel of the output has a correspondent in the input (i.e. no epenthetic vowels).

Further, I make use of constraints regulating syllabic well-formedness, specifically the ONSET (ONS) and NOCODA constraints listed in (5):

- (5) Syllabic Harmony Constraints:
 ONS Syllables must have onsets.
 NOCODA Syllables are open.

Interior VLE has a minimal word requirement, specifically that the Prosodic Word (PrWd) be at least two syllables in length. Following Downing (1997), I make use of the MINIMALITY (MIN) constraint set out in (6):⁴

- (6) MIN Prosodic Words (PrWd) can be no smaller than two syllables.

The first point in looking at how these constraints bear upon Interior VLE is that, even though MIN requires that words be minimally disyllabic, when the input has the shape CV, other constraints outrank MIN, specifically DEP-C and ONS, and the optimal candidate is monosyllabic, as illustrated by the tableau in (7).⁵

(7) Interior VLE /du/ ‘to do’

	DEP-C	ONS	MIN	DEP-V	MAX
¹ du			*		
du.i		*!		*	
du.ti	*!			*	

Two of the candidates in (7) display a paragogic vowel; however, the optimal candidate — which violates neither DEP-C nor ONS — is one that displays no paragogic vowel.


In contrast to (7), when the input has the shape CVC, a candidate containing a paragogic vowel is selected as optimal. In this case, MIN is not violated. The tableau in (8) illustrates this, with *te.ke* emerging as the optimal candidate for /tek/.

(8) Interior VLE /tek/ ‘take’

	NOCODA	MIN	DEP-V	MAX
¹ te.ke			*	
te		*!		
tek	*!	*!		

When the input is already disyllabic, as in (9), then a word-final consonant is not necessary for the construction of a minimal PrWd and does not appear in the optimal output. Thus *rɛ.spɛ* is the optimal candidate in (9), not **rɛ.spɛ.ke* or **rɛ.spɛk*.

(9) Interior VLE /rɛspɛ/ ‘respect’

	NOCODA	MIN	DEP-V	MAX
 rɛ.spɛ				*
rɛ.spɛ.ke			*!	
rɛ.spɛk	*!			
rɛs.pɛ	*!			*

A further point with regard to Interior VLE verbs is that the “two-word” verbs of English, e.g. *call up*, behave as single units in VLE. This is true syntactically, semantically, and phonologically. It is illustrated in (10), where the pronominal object obligatorily comes after the particle rather than before it as it would in English:

(10) Interior VLE

A mə tek a dɛ?
 I must take out them
 ‘Should I take them out?’

Consequently, it is appropriate to include both the verb and the particle within a single PrWd. Thus, *te.ka* in (10) satisfies MIN.⁶

The constraints and constraint rankings presented thus far account for the phonological shape of verbs in Interior VLE. A summary of the rankings is given in (11). The leftmost constraints there, i.e. DEP-C and NOCODA, are both unviolated. Forms like *pe a*, i.e. ‘pay out,’ establish that DEP-C is more highly ranked than ONS.

(11) Constraint ranking in Interior VLE:

DEP-C >> ONS >> MIN >> DEP-V >> MAX
 NOCODA >> DEP-V >> MAX

I asserted above that substratal phonology can play a major role in shaping PC phonology by the transfer of the substrate’s constraint rankings to the PC. Table 1 shows that the constraints that are highly ranked in Interior VLE are ones that are highly ranked in Loma as well.

Table 1. A comparison of Interior VLE with Loma

Highly ranked VLE constraints	Their status in Loma (Sadler 1951; Dwyer 1981)
NOCODA	Inviolable in Loma
DEP	Inviolable in Loma
MIN	MIN is highly ranked in Loma. The only forms that violate it are the few cases where the input is /CV/. Violated in Loma but not frequently. All PrWd's begin with a consonant.
ONS	Violated in Loma but not frequently. All PrWd's begin with a consonant.
MAX	According to Sadler (1951:314-6), a glide deletes in some environments when the vowel that follows it agrees with it for backness; a voiced velar fricative is sometimes deleted when it precedes <i>a</i> . These are the only reported violations of MAX.

As far as coda consonants are concerned, the ranking of two constraints in particular accounts for the differences between Loma and *non*-pidginized English. In Loma, NOCODA is inviolable and MIN, while not inviolable, is highly ranked. In non-pidginized English, on the other hand, both NOCODA and the disyllabic application of MIN are ranked low. In its ranking of these constraints, Interior VLE patterns with Loma, not with non-pidginized English.

The discussion thus far has been limited to verbs. It is appropriate to look at Interior VLE more generally, particularly with reference to MIN. In Loma, MIN applies across the board; words must be disyllabic regardless of their lexical category. In contrast, in Interior VLE, paragogic vowels occur with verbs but not with words in other lexical categories. The data in (12) illustrate this.⁷

(12) Interior VLE

<i>verbs</i>		<i>non-verbs</i>
mi.ti	'to meet'	mi 'meat'
flo.ge	'to flog'	fro 'frog'
lo.de	'to load'	ro 'road'

At first glance, the data presented in (12) suggest that MIN is highly ranked in Interior VLE only with reference to verbs. One way to capture this would be to adjust MIN, recasting it as in (6').

(6') MIN-VERB A verb can be no smaller than two syllables.

Invoking (6') would yield the appropriate tableaux in Interior VLE. But a

disjointed account is surely less desirable than a unified one. Further, positing MIN-VERB for VLE obscures VLE's link to Loma, where MIN is highly ranked regardless of a word's part of speech. I wish to argue that MIN-VERB is in fact unnecessary. Rather, just as MIN is highly ranked in Loma, so it is highly ranked in Interior VLE, i.e. MIN is category-insensitive in Interior VLE in the same way that it is category-insensitive in Loma. To understand the basis for this claim, it is necessary to look at the broader history of English-lexifier pidgins in Liberia.

5. Word-final Consonants in Coastal VLE

Attestations of the pidgin along the Liberian coast begin in the eighteenth century. As discussed in Singler (1997), pidgin was widely spoken along the Liberian coast even before the African-American Settlers started arriving there in the 1820's. Once the Settlers arrived and seized control of the region, their English came to serve as the superstrate language for this existing pidgin. Today that pidgin, VLE, continues to be a variety distinct from Settler English.



Map 1. The distribution of VLE's primary substratal languages

A consideration about VLE is that throughout the nineteenth century it, like Settler English, was strictly a coastal phenomenon. Then, at the beginning of the twentieth century, the Liberian government set out to exert control over the regions interior to the coast. This was a slow process but, by the middle part of the century, VLE had been disseminated throughout the Liberian interior.

Because VLE evolved over a period of one to two centuries on the coast before it spread into the interior, it becomes appropriate to look at Coastal VLE and the substrate languages influencing it. The most important substrate languages for Coastal VLE are Kru languages, specifically Bassa and Klao. The distribution of substratal languages is given in Map One. The languages on the coast in the map are Kru languages; those in the interior are Mande.⁸

Like the interior Mande languages, Bassa and Klao permit no coda consonants. However, MIN is ranked very low. Kru languages do not require that PrWd's be minimally disyllabic.

The ranking of the constraints in Bassa and Klao is stated in (13). The same ranking holds for Coastal VLE.

(13) NOCODA >> DEP >> MIN, MAX

A consequence of the ranking in (13) is given in (14), where *te*, not **te.ke*, is the optimal Coastal VLE form.

(14) Coastal VLE /tek/ 'take'

	NOCODA	DEP	MIN	MAX
te			*	*
tek	*!		*	
te.ke		*!		

As (14) illustrates, in Coastal VLE the optimal response to English coda consonants is to eliminate them from the output. This is true regardless of the word's lexical category, as shown in (15):

(15) Coastal VLE

verbs

mi 'to meet'

flɔ 'to flog'

lo 'to load'

non-verbs

mi 'meat'

frɔ 'frog'

ro 'road'

Why is there a difference between Coastal and Interior verbs (and not between Coastal and Interior VLE non-verbs)? The answer to that lies in the fact that

even though the optimal candidate in Coastal VLE for ‘take’ is *te* and for ‘reach’ is *ri*, the final consonants of these verbs do show up elsewhere in the language. For, while it is characteristic of PCs to have very little morphology and while this holds true of Coastal VLE, the verb suffix *-ẽ* (< Eng. *-ing*) is present. As a result, the word-final coda consonant in English verbs continues to appear in Coastal VLE output, in the *-ẽ* form if not the bare stem form, as (16) shows:

(16) Coastal VLE

<i>verbs</i>		<i>suffixed verbs</i>		<i>non-verbs</i>	
mi	‘to meet’	mi.tẽ	‘meeting’	mi	‘meat’
flo	‘to flog’	flo.gẽ	‘flogging’	frɔ	‘frog’
lo	‘to load’	lo.dẽ	‘loading’	ro	‘road’

Note that nouns do not display comparable paradigms. Number is ordinarily marked by the postposed consonant-initial morpheme *den*, and possession is marked strictly by word order. Thus, while there is an alternation in verb forms in which a coda consonant of the stem can become the onset consonant in a suffixed bimorphemic form, no such alternation exists for other parts of speech. As a consequence, the coda consonants of English verbs show up in certain contexts in the output of Coastal VLE, while the coda consonants of English nouns and adjectives and other parts of speech never do so.

The distribution of English coda consonants in Coastal VLE indicates that, while English itself provided the input for Coastal VLE, it is Coastal VLE that provided the input for Interior VLE. And because Coastal VLE — rather than English directly — provided the basis for the input of Interior VLE, verbs are the only words whose inputs have coda consonants. No word class other than the verb retains the final C in a CVC string in Coastal VLE; hence no word class other than the verb has a coda consonant in Interior VLE input. In (17), I repeat the Interior VLE data from (12), but this time I provide the input as well:

(17) Interior VLE

<i>verbs</i>			<i>non-verbs</i>	
/mit/	mi.ti	‘to meet’	/mi/	mi ‘meat’
/floŋ/	flo.ge	‘to flog’	/frɔ/	frɔ ‘frog’
/lod/	lo.de	‘to load’	/ro/	ro ‘road’

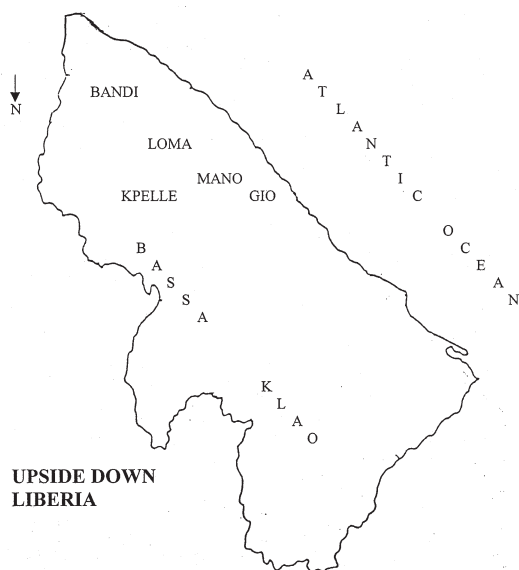
It can now be seen that the MIN-VERB constraint proposed in (6') is unnecessary. VLE makes use of MIN in a category-insensitive way. While it is a

consequence of the history of Interior VLE that verbs are the only words that have the shape CVC, that is a fact expressed in the input of lexical items rather than by making use of a category-sensitive constraint.

The data that I have presented make a simple point: OT allows for a more straightforward characterization of the nature of substratal influence on a PC. In crucial ways, Coastal VLE inherited its constraint rankings from its Kru substrate. The coastal VLE output then formed the basis for the input to Interior VLE, a variety which inherited its constraint rankings from its Mande substrate.

6. “What if?”

Having shown this, I would like to indulge in a “what if?” scenario. Specifically, I want to ask “what if,” say, 500 years ago, Liberia had been turned upside down, so that the speakers of Loma and other relevant Mande languages lived along the coast while the speakers of Bassa and Klao lived solely in the interior, *à la* Map Two.



Map 2. The distribution of VLE's primary substratal languages in an “Upside-down” Liberia

If the Mande speakers had been the first to acquire VLE, then the input would have come from English directly, rather than from the output of an older VLE that had been influenced by Kru phototactics. In this new scenario, the high ranking of MIN (derived from Mande) would have preserved the coda consonants of the English input for all parts of speech, as in (18):

- (18) “Upside-down” Coastal VLE (spoken by Mande speakers)
- | | | | |
|-------|-----------|-----------|--------|
| verbs | | non-verbs | |
| mi.ti | ‘to meet’ | mi.ti | ‘meat’ |
| fɔ.ge | ‘to flog’ | fɔ.ge | ‘frog’ |
| lo.de | ‘to load’ | ro.de | ‘road’ |

Assuming the same historical sequencing, this would now mean that “Upside-down” Interior VLE, i.e. the variety spoken by Kru speakers, would have based its input on surface forms from “Upside-down” Coastal VLE. Assuming the “Upside-down” Interior VLE constraint rankings to be essentially those of its speakers’ first languages, i.e. Bassa and Klao, faithfulness to the input would have prevailed, as shown in (19):

- (19) “Upside-down” Interior VLE (spoken by Kru speakers)
- | | | | |
|-------|-----------|-----------|--------|
| verbs | | non-verbs | |
| mi.ti | ‘to meet’ | mi.ti | ‘meat’ |
| fɔ.ge | ‘to flog’ | fɔ.ge | ‘frog’ |
| lo.de | ‘to load’ | ro.de | ‘road’ |

A comparison of actual VLE to “Upside-down” VLE is given in Table 2:

Table 2. A comparison of Real World VLE with Upside Down VLE

a. Actual first languages:	Coastal VLE - Kru		Interior VLE - Mande		
	/mit/	[mi]	/mit/	[mi.ti]	‘to meet’
	/mitẽ/	[mi.tẽ]	/mitẽ/	[mi.tẽ]	‘meeting’
	/mit/	[mi]	/mi/	[mi]	‘meat’
b. “Upside-down” first languages:	Coastal VLE - Mande		Interior VLE - Kru		
	/mit/	[mi.ti]	/miti/	[mi.ti]	‘to meet’
	/mitẽ/	[mi.tẽ]	/mitẽ/	[mi.tẽ]	‘meeting’
	/mit/	[mi.ti]	/miti/	[mi.ti]	‘meat’

My point in playing “what if?” is this: Why are the VLE forms that actually occur the ones that are in the top half of Table 2 and not the ones in the bottom half? The answer, I argue, is a direct consequence of the sequence in which the pidgin came into contact with specific substrate languages. In the real world, VLE came into contact with both Kru and Mande languages, but there was a particular order to the contact: the Kru languages first, the Mande languages later. It is my contention that, had this order of interaction been reversed, some surface forms in both the Coastal and the Interior VLE would have been different.

A crucial difference between Atlantic pidgins and Atlantic creoles is that pidgins do not ordinarily involve massive displacement of populations, while creoles do. That means that one often has a greater degree of confidence in speaking of specific substrate languages and their influence in the case of pidgins than in that of creoles. In the Liberian case, when it comes to substratal influence on the pidgin, geography has dictated chronology. In the Caribbean case, while there is no neat correlation between geography and chronology, there is a chronological component to substratal input nonetheless. Singler (1996b), for example, argues that the Antillean French-lexifier creoles differ from one another in part because different African language groups predominated at different points in their histories. The evidence from word-final consonants in VLE shows still more support for the assertion that it does not suffice to know solely where competing substratal influences came from; one must also know the sequence in which those influences were present, because a different sequence can yield a different outcome. While the evidence I have offered is phonological, the cruciality of historical sequencing surely applies to an entire creole grammar.

Notes

1. I am grateful to Darlene LaCharité, John McWhorter, Katya Zubritskaya, An-Nah Moon, and Keith Fernandes for their helpful suggestions. All errors are my own. An NYU Research Challenge Fund Grant made possible the research upon which this paper is based. I am grateful to the elders of Borkeza for their willingness to be interviewed and to David Peewee, Boakai Zoludua, and especially Sumoyea Guluma for their invaluable assistance.
2. In actual speech as opposed to idealized speech, particularly when the actual speech comes from what is asserted to be the basilectal range of a PC, there are — perhaps inevitably — often non-basilectal elements that show up in the data. In (1), that is the case

with the three coda consonants, i.e. the final *m*'s of *Enitam* and *fam* and the final *s* of *es*. Inasmuch as predicate adjectives do not take an overt copula in the basilect, the very presence of *es* is itself non-basilectal. The example of Interior VLE in (2) below is like the example of Coastal VLE in (1) in having non-basilectal elements present, specifically the coda consonants of *das* and the noun *wɔk*.

While I assume the source of *chakla*, the final word in (1), to be Niger-Congo rather than English, I do not know more about the source of its origin.

3. On the basis of the Borkes corpus, Singler (1996a) makes the case for positing two basilects in Interior VLE, the Alpha Basilect being further from the acrolect than is the Beta Basilect; that article then relates these two basilects one to the other. The speech variety discussed in the present study is the Alpha Basilect.
4. Mühlhäusler (1986:150) notes a tendency in the pidgins of the world for lexical items to have a disyllabic word structure. However, Mühlhäusler's comment is meant to point out that words in a pidgin have fewer syllables on average than words in its lexifier language. Thus, while Mühlhäusler's concern is with the average length of pidgin lexical items, mine is with the minimum permissible length of such words in Interior VLE.
5. The primary metric for formalizing OT is the tableau, of which (7) is an example. The candidates in the left column represent potential output; the constraints are ranked from left to right in decreasing order of importance. When there is a dotted line between two constraints, this indicates that their ranking relative to each other is unresolved or indeterminate. (This is illustrated below in [8], where NOCODA and MIN are unranked relative to each other.) Candidates are compared by evaluating their violation (or non-violation) of constraints. That candidate is optimal which is "last" to violate a constraint as one moves from left to right in the tableau (or which violates none of the relevant constraints whatsoever). However, if all candidates violate a given constraint an equal number of times, then the constraint does not act to eliminate candidates. If all candidates violate the same constraint but some violate the constraint more frequently than others, then those candidates that violate it more than the minimum number of times are knocked out. The exclamation mark signals a "fatal" violation, and the hand with the pointed finger next to a candidate indicates that it is the optimal candidate.
6. Singler (1996a) discusses additional constraints that bear upon the present data set, most of them pertaining to possible onsets. For example, resyllabification is not possible when its occurrence results in an impermissible onset cluster. Singler (1996a) also acknowledges a complication that must be addressed: the Borkes corpus, while showing a general avoidance of consonants in the coda position in the output, does appear to allow bilabial stops there. That article draws on work by Kiparsky (1993, 1994) to provide a solution within OT, but the possibility remains that there is a skew in the data, one that will disappear when a more extensive corpus is utilized.
7. There is some variation between *flɔ.ge* and *flɔ.ke*. While that variation does not bear directly on the analysis of the minimal word and will not be considered in this article, its existence is consistent with the historical scenario discussed below, specifically with the fact that the Kru language Klao has /k/ but not /g/ in its phonemic inventory.
8. There is a lone Mande language spoken on the Liberian coast, Vai, spoken near the border with Sierra Leone. Geography, history, and demography (there are far fewer speakers of Vai in Liberia than of Bassa or of Klao) have conspired to restrict the role of Vai in the development of VLE.

References

- Archangeli, D. 1997. "Optimality Theory: An Introduction to Linguistics in the 1990s". In *Optimality Theory: An Overview*, D. Archangeli and D. T. Langendoen (eds.), 1-32. Malden, MA: Blackwell.
- Arends, J. 1986. "Genesis and Development of the Equative Copula in Sranan". In *Substrata Versus Universals in Creole Genesis*, P. Muysken and N. Smith (eds.), 103-27. Amsterdam: John Benjamins.
- Chaudenson, R. 1979. *Les créoles français*. Paris: Fernand Nathan.
- Chaudenson, R. 1992. *Des îles, des hommes, des langues*. Paris: L'Harmattan.
- Chomsky, N. and M. Halle. 1968. *The Sound Patterns of English*. New York: Harper and Row.
- DeCamp, D. 1971. "Toward a Generative Analysis of a Post-Creole Speech Community". In *Pidginization and Creolization of Languages*, D. Hymes (ed), 349-70. Cambridge: Cambridge University Press.
- Downing, L. J. 1997. "Prosodic Stem \neq Prosodic Word in Bantu". Paper presented at the Berlin Conference of the Phonological Word.
- Dwyer, D. J. 1981. *Lorma: A Reference Handbook of Phonetics, Grammar, Lexicon and Learning Procedures*. East Lansing: Michigan State University, African Studies Center.
- Kiparsky, P. 1993. "Variable Rules". Rutgers Optimality Workshop.
- . 1994. "Remarks on Markedness". Trilateral Phonology Weekend II, UC Santa Cruz.
- McCarthy, J. and A. Prince. 1993a. "Generalized Alignment". In *Yearbook of Morphology 1993*, G. Booij and J. van Marle (eds.), 79-153. Dordrecht: Kluwer.
- . to appear. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Cambridge, MA: MIT Press.
- . 1994. "The Emergence of the Unmarked". In *NELS 24: Proceedings of the 1993 Meeting of the Northeastern Linguistic Society*, M. González (ed.), 333-79. Amherst, MA: Graduate Student Linguistic Association.
- . 1995. "Faithfulness and Reduplicative Identity". In *Papers in Optimality Theory*, J. Beckman, S. Urbanczyk, and L. Walsh (eds.), 249-384. Amherst, MA: Graduate Student Linguistic Association.
- Moon, A. 1996. Aspects of Old English Prosody: An Optimality-Theoretic analysis. New York University Ph.D. dissertation.
- Mühlhäusler, P. 1986. *Pidgin and Creole Linguistics*. Oxford: Basil Blackwell.
- Prince, A. and P. Smolensky. 1993. "Optimality Theory". Rutgers University / University of Colorado, Boulder MS.
- Sadler, W. 1951. *Untangled Loma*. Monrovia, Liberia: Board of Foreign Missions of the United Lutheran Church in America for the Evangelical Lutheran Church in Liberia.
- Singler, J.V. 1984. Variation in Tense-Aspect-Modality in Liberian English. University of California, Los Angeles Ph.D. dissertation.
- . 1986. "Short Note". *Journal of Pidgin and Creole Languages* 1:141-45.
- . 1988. "The Place of Variation in the Formal Expression of Inflectional Processes: Evidence from Kru Pidgin English". In *Linguistic Change and Contact: Proceedings of the Sixteenth Annual Conference on New Ways of Analyzing Variation in Language*, K.

- Ferrara, B. Brown, K. Walters, and J. Baugh (eds.), 345-53. Austin: University of Texas Department of Linguistics.
- . 1989. "Plural Marking in Liberian Settler English, 1820-1980". *American Speech* 64: 40-64.
- . 1990. "The Impact of Decreolization upon TMA: Tenselessness, Mood, and Aspect in Kru Pidgin English". In *Pidgin and Creole Tense-Mood-Aspect Systems*, J. V. Singler (ed), 203-30. Amsterdam: John Benjamins.
- . 1991. "Phonology in the Basilect: The Fate of Final Consonants in Liberian English". *Studies in African Linguistics* 22:1-44.
- . 1996a. "An OT Account of Pidgin Phonology: Coda consonants in Vernacular Liberian English". In *Proceedings of the Twenty-second Annual Meeting of the Berkeley Linguistics Society*, J. Johnson, M. L. Juge, and J. L. Moxley (eds.), 375-86. Berkeley: Berkeley Linguistics Society.
- . 1996b. "Theories of Creole Genesis, Sociohistorical Considerations, and the Evaluation of Evidence: The case of Haitian Creole and the Relexification Hypothesis". *Journal of Pidgin and Creole Languages* 11:185-230.
- . 1997. "The Configuration of Liberia's Englishes". *World Englishes* 16: 205-31.
- . forthcoming. "Assessing the Legitimacy of Enclave Evidence in Reconstructing the History of African American English: The Liberian Settler English of Sinoe". In *African American English and its Congeners*, John M. Lipski (ed). Amsterdam: John Benjamins.
- Weinberg, J. P. 1996. "Past Tense Verb Reduplication in Luiseño". Paper presented at the annual meeting of the Linguistics Society of America, San Diego.

The Story of *kom* in Nigerian Pidgin English

Sali A. Tagliamonte*

1. Introduction

Recent years have seen an increasing number of studies of grammaticalization in contact languages (e.g. Baker & Syea 1996). Due to the extreme socio-historical circumstances under which pidgins and creoles have typically developed, normally gradual processes of linguistic restructuring are compressed, giving the advantage of “telescoped diachronic evidence”, while historical and ongoing developments in the source languages may often provide the backdrop for the interpretation of concurrent synchronic variation (Mufwene 1996a:25). Such circumstances provide a timely opportunity to examine the underlying mechanisms of linguistic change.

However, the precise role of grammaticalization in creolization remains controversial, as ongoing research suggests that pidgin and creole genesis contexts give rise to a number of different *types* of grammaticalization, some of which are apparent, rather than authentic (Bruyn 1995; see also Plag 1998). The crucial issue is whether or not the linguistic processes observed in pidgin and creole language situations mirror actual diachronic developments.

On the other hand, the growing literature on grammaticalization processes more generally, in which linguistic categories are conceived as the result of a series of transitions forming a path, or trajectory over time (Hopper & Traugott 1993), demonstrates that synchronic analysis can reveal form-function asymmetries which relate to the stage of development of the process under investigation. Moreover, a number of recurring grammatical “principles” can be used to identify and measure the precise nature and extent of grammaticalization, specifically the strength and significance of multiple independent linguistic factors acting upon the evolving grammatical mor-

pHEME (Poplack & Tagliamonte 1995, 1996a, 1996b). Such correlations are crucial to examining the mechanism of diachronic grammaticalization in synchronic data (Traugott & Heine 1991). With respect to the restructuring processes themselves, some researchers suggest that their origins may be found in discourse-level phenomena (Traugott 1982), particularly in pidgin and creole situations (Givón 1979a, 1984). Indeed, quantitative empirical analyses have provided evidence to suggest that the grammatical functions of some markers may be the logical outgrowth of earlier, more facultative discourse-level phenomena (Sankoff & Brown 1976, Sankoff 1990).

In this study, I examine the pre-verbal marker *kom* in Nigerian Pidgin English (NPE), an expanded pidgin as well as a variety which is said to be rapidly nativizing (Agheyisi 1984, Faraclas 1987, Singler 1992). Taking the entire past temporal reference sector as a point of departure, my approach is innovative in examining both narrative and information structure factors and the correlation of *kom* with discourse markers alongside traditional syntactic/semantic factors in order to identify its function. It corroborates quantitative research which has demonstrated that there is no absolute unmarking of verbs in complicating action clauses in a number of creole languages (Myhill 1991, Patrick 1999, Rickford 1987a), as had been predicted in earlier (non-quantitative) studies (Corne 1973, Bollée 1977, Givón 1979a). In order to track the development of this feature in apparent time I also provide a time-depth perspective through comparison across two recent generations of NPE speakers.

2. NPE *kom*

Kom is perhaps the most productive and widespread tense/aspect feature of the NPE data under consideration.¹ The examples (1)-(7) illustrate the breadth of usage of this form. While it is clearly situated in PAST time, it can modify both stative and non-stative verbs, as in (1) and (2):

- (1) a. *A **kɔm** deh respet dat wɔn.* (5/373)
'I was respecting that one'
- b. *I **kɔm** no ma senyɔ brɔda.* (13/192)
'He knew my older brother.'
- (2) a. *Ma hɔsban **kɔm** sen mi di pepas.* (13/112)
'My husband sent me the papers'

- b. *I kɔm open dɔ fɔ mi.* (13/047)
 ‘He opened the door for me’

It often co-occurs with the progressive/iterative marker *deh*, as in (3), and sometimes in copula constructions with *dey*, as in (4). Combinations with other pre-verbal markers, e.g. *don*, are infrequent:

- (3) a. *So wi kɔm deh draiv deh go.* (13/035)
 ‘So we were driving on’
 b. *Get wɔn dey wey rain kɔm deh fɔl.* (4/224)
 ‘One day rain was falling.’
- (4) a. *Ma lugaɲ no kɔm deh de.* (13/14)
 ‘My luggage wasn’t there.’
 b. *Ma fotograf no kɔm de fɔ dɔ paspɔt.* (4/120)
 ‘My photograph wasn’t in the passport.’

Kom may also appear in contexts where it may still be interpreted literally, as in (5):

- (5) a. *De kɔm bak.* (4/14.55)
 ‘They came back.’
 b. *De mama kɔm fɔ ma mama haus.* (4/1.49)
 ‘Their mother came to my mother’s house.’
 c. *Yu no se mi ay ʃus kɔm frɔm Niʒa nao.* (13/5.12)
 ‘You know that I just came from Nigeria.’
 d. *A kɔm frɔm hom wey de disIplɪn us well well.* (13/29.08)
 ‘I came from a home where they disciplined us very well.’

However, the fact that the two functions co-occur, as in (6), suggests that *kom* has an independent auxiliary function:

- (6) a. *Wen i bai sɔmtɪŋ nao, na im i kɔm kɔm haus.* (4/25-6)
 ‘Then when he bought something, he came to the house.’
 b. *Wen im kɔm kɔm agen, e no kɔm si eni snayl.* (4/27.16)
 ‘When he came again, he didn’t see any snail.’
 c. *So, di paspɔt kɔm kɔm.* (13/12.29)
 ‘So, the passport came.’
 d. *Evribodi kɔm kɔm back to hi own haus nao, i own tawn.*
 (13/20.04)
 ‘Everybody came back to his own house now, his own town.’

Other uses of *kom* reveal that the auxiliary form may still retain some of its literal meaning, as in, (7a), particularly in contexts where it forms part of a serial verb construction, as in (7b), or combines with other markers, as in (7c), making a number of readings possible:

- (7) a. *Evriwea kom be so so banana tri.* (4/017)
 ‘Banana trees were everywhere.’
 ‘Banana trees came to be everywhere.’
- b. *I kom sIdaon fɔ buʂ de.* (4/012)
 ‘He sat down in the bush there.’
 ‘He came and sat down in the bush there.’
- c. *Winta kom deh kom.* (2/130)
 ‘Winter was coming.’
 ‘Winter kept on coming.’

Such contexts reveal synchronic diversity. Indeed, as is well-known, the changes associated with grammaticalization are actually a continuum. During any phase of development, co-existence of a form which is undergoing grammaticalization will compete with other constructions, some of which may involve clear pragmatic differences, but others which will be quite similar in function (Hopper 1991:22).

What is the status of *kom* in these NPE data? Poplack & Tagliamonte (1996b) argue that *kom* is a central part of NPE’s relative tense system serving to denote sequential temporal relationship. On the other hand, Faraclas (1987) claims that auxiliary *kom*, (or *ka* as he refers to it), is “a marker of objectivity or realis modality” (ibid. 50) which has no direct bearing on time reference or sequence. Moreover, Faraclas claims that the use of *kom* has shifted dramatically between two successive generations of NPE speakers, being found with increasing frequency amongst Nigerian youth, having grammaticalized according to “universal patterns of language change”.

Although providing opposing views with respect to the function of *kom*, both studies concur that it is a grammaticalizing form. However, its precise characterization, stage of development, and particularly the trajectory of its development remain unknown. Thus, it is important to determine the relative frequency of forms at any particular point in time in order to determine which one of them predominates. Their distribution according to independent features of grammar is especially critical in order to track the process by which the potentially grammaticalizing form may be extending its range of meanings

in its pathway of change.

Moreover, as pointed out by Sankoff (1990:309) in connection with *ben* in Sranan and Tok Pisin, a typical pattern in the initial stages of grammaticalization in pidgin and creole languages is that selection of an overt form is not an automatic part of the syntax, but proceeds according to an optional system. She suggests that while semantic distinctions are being created, use of the forms expressing them will be sensitive to discourse-pragmatic contexts.

In order to assess such questions two types of evidence are necessary: 1) data from two different points in time, and 2) a multi-level analysis. In the ensuing analyses I contribute to these issues by providing a detailed account of the function of *kom* from a syntactic/semantic and discourse-pragmatic perspective in speakers from recent successive generations. What patterns of change may be observed in the development of *kom* across these speakers which might elucidate the pathway of grammaticalization? If change is observed, what is the underlying mechanism by which the change has taken place?

3. Data

The study is based on data collected by Ejike Eze in the summer of 1994 through participant observation in his own social networks in the Nigerian community in Ottawa, Canada (Tagliamonte, Poplack & Eze 1997). In this paper I focus on the use of *kom* in an adult and child pair from this corpus, matched for their socio-economic profile, linguistic background and the density of their social network links. The child is a girl of eleven who was interviewed only weeks after arriving in Canada from Nigeria. The adult is a 35-year-old woman who had lived in Canada for about a year at the time of the interview. As the analyses will reveal, the adult and child exhibit systematic use of *kom* which is virtually identical, with only hints that the child's language shows a more advanced stage of grammaticalization of *kom*.

4. Distributional Analysis

First, I situate the similarities and/or differences between adult and child speakers within the more general past temporal reference sector of the grammar. Table 1 compares the overall distribution of variants used by each

speaker in past temporal reference contexts based on earlier work on these materials (Poplack & Tagliamonte 1996b).

Unlike other empirical studies of pidgin and creole varieties, where the vast majority of verb forms are typically found to be stems, here a full 52% of the verbs used by the child and 44% of those used by the adult are marked with *kom*. Use of other markers like *deh* or *bin* or *dey* are relatively infrequent, accounting (in combination) for 11% and 21% of the total past temporal reference contexts, respectively. However, the most striking finding of Table 1 is that in contrast to what might be expected of a feature undergoing rapid change, there is very little to distinguish between the two generations, particularly in the two most productive forms, the verb stem and *kom*.

It might be argued that these two speakers represent some acrolect-influenced variety, possibly from Nigerian Standard English, or even the variety of English spoken in Ottawa, hence the similarities. However, recall

Table 1. Overall distribution of verb forms in PAST TEMPORAL REFERENCE contexts in NPE across two generations

VARIANTS:	CHILD		ADULT	
	%	N	%	N
TOTAL KOM	52	183	44	241
<i>kom</i> + verb		139		190
<i>kom</i> + <i>deh</i> + verb		36		34
<i>kom</i> + <i>dey</i> + verb		2		7
<i>kom</i> + <i>be</i>		1		2
<i>kom</i> + verb + <i>finiš</i>		5		8
VERB STEM:	36	127	35	192
TOTAL DON	3	9	8	45
<i>dɔn</i> + verb		6		36
<i>dɔn</i> + verb+ <i>finiš</i>		1		4
<i>dɔn</i> + <i>deh</i> + verb		1		3
<i>dɔn</i> + <i>deh</i>		Ø		1
<i>dɔn</i> + <i>be</i>		1		1
DEH: <i>deh</i> + verb	5	17	8	45
BIN: <i>bin</i> + verb		Ø		1
Locative copula		6		17
English-type marking		Ø		1
<i>fɪnɪʃ</i> + verb		7		2
Existential copula		Ø		3
TOTAL		349		547

that the child had only been in Canada a few weeks. Moreover, the frequency of use of *kom* by the adult is comparable to the child's. Finally, the overall frequency of *kom* by both speakers is actually far *in excess* of that reported in Faraclas (1987) for an eastern dialect of NPE collected in 1985 (compare frequencies of approximately 30% in L1 speakers, to 44% for the adult and 52% for the child in the current sample). Thus, there does not seem to be any indication, at least from this vantage point, for interference from English. More importantly however, there does not appear to be any discernable difference in their use of past temporal reference markers.

4.1. Syntactic/semantic features

Frequency alone, however, provides only a gross perspective on the data. In order to determine the *function* of *kom*, and whether it is being used for the same or different function(s) by both speakers, I now examine its distribution.

The series of graphs that follow present a factor-by-factor analysis of the patterns of variability across a number of well-known and well-studied indices of tense/aspect morphology in pidgin and creole languages.

Figure 1 shows the distribution of *kom* according to the STATIVE/NON-STATIVE distinction.

The figure shows that there is a clear difference between the occurrence of *kom* on stative vs. non-stative verbs, with non-statives appearing with *kom* at more than double the rate of statives. This trend is parallel in direction and degree in adult and child.

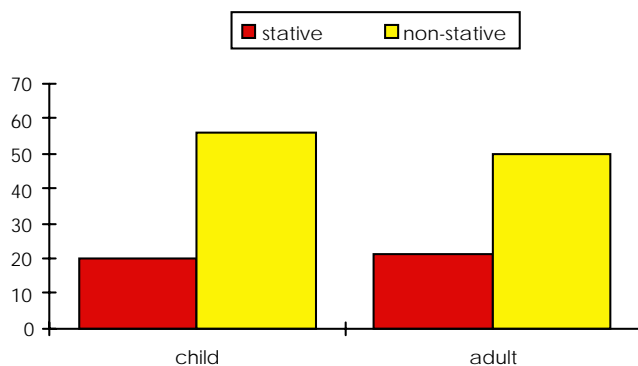


Figure 1. Distribution of *kom* across stative and non-stative contexts.

Figure 2 shows the distribution of *kom* according to temporal relationship. Here sequential temporal relationships show the greatest concentration of *kom* followed by non-anterior and finally anterior temporal relationships. Again, the pattern, both in direction of effect and degree are parallel between adult and child.

Figure 3 shows the distribution of *kom* across negative and affirmative sentence types. *Kom* appears much more frequently in affirmative contexts than negative contexts. Once again, the pattern is the same between adult and child.

These results confirm the received wisdom about NPE *kom*, that it is sensitive to the STATIVE/NON-STATIVE distinction and is part of a RELATIVE TENSE system. There also appears to be an interdiction against overt marking in NEGATIVE sentences with past temporal reference.² However, for *kom* — at least with respect to these factors — there is no evidence of change or restructuring in apparent time. This provides additional support for the fact that the two speakers share the same system, at least with respect to past temporal reference.

While tabulations of effects taken one at a time, such as those reported above, are informative, they do not reveal the *relative* importance of factor effects to each other, nor whether all are in fact significant when considered simultaneously. Moreover, this view of the data is primarily limited to *syntactic/semantic* features related to the TENSE/ASPECT organization of the grammar, and does not consider the contextual clues and discourse-pragmatic features to which grammaticalizing categories are reputed to be sensitive.

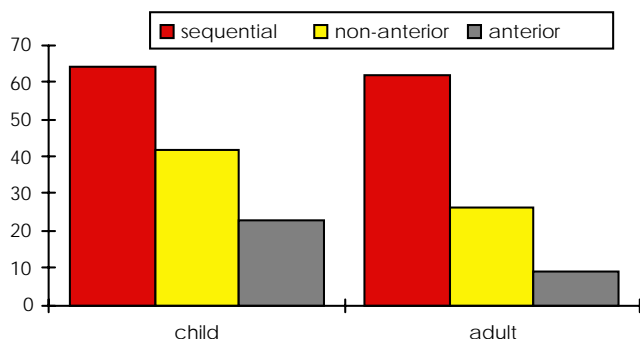


Figure 2. Distribution of *kom* across temporal relationship contexts.

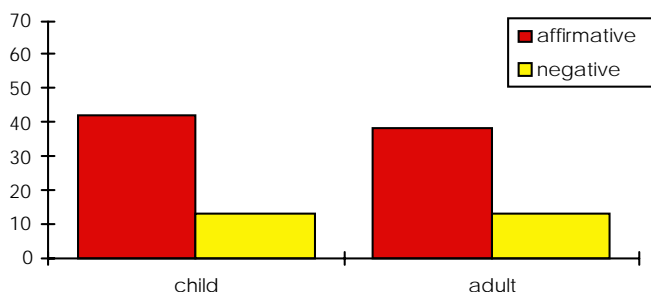


Figure 3: Distribution of *kom* across negative and affirmative contexts.

4.2. Discourse/pragmatic features

As mentioned earlier, researchers who have explored syntactic change in pidgin and creole contexts suggest that linguistic changes may actually originate in discourse phenomena (Sankoff & Brown 1976, Givón 1979a, Traugott 1982, Sankoff 1990). For example, the changes undergone by a grammatical element may be observed in the way it is utilized in the organization of the discourse as it shifts from one type of grammatical function to another. If *kom* is a linguistic feature currently undergoing grammaticalization, or a marker that completed its grammaticalization path recently, then these factors may well be critical in elucidating this pathway of change. For example, we might anticipate not being able to distinguish between discourse-pragmatic features and syntactic/semantic features, or indeed a greater dependence on the latter in the process of linguistic change.

Accordingly, all the verbs in the data were also categorized according to a number of discourse-pragmatic features, including PRESENCE OF A DISAMBIGUATING TEMPORAL ADVERB and MARK ON THE PRECEDING REFERENCE VERB — to test general predictions about the role of contextual disambiguation in marking variability in creoles — and the FOREGROUND/BACKGROUND distinction in discourse discussed by Hopper (1979) and Givón (1979a).

The data were then re-analyzed using the logistic regression procedures implemented in the variable rule program (Rand & Sankoff 1990) to estimate the true effects of all these factors when all of them are considered simultaneously, as well as to remove artifacts of poor data distribution, correlated

factors, or statistical fluctuations typical of language data. This method permits assessment of which factors have a significant effect on *kom*, as well as an estimation of the magnitude of individual factor effects. Finally, a comparative analysis between child and adult will permit assessment of the nature of similarities and/or differences between the two generations, and thus an assessment of the presence and nature of any putative grammatical change.

Table 2 depicts two multivariate analyses of the contribution of factors selected as significant to the probability of *kom* in past temporal reference contexts.³

Surprisingly, Table 2 shows that only two factors are statistically significant to the occurrence of *kom*: — SEQUENTIAL temporal relationships and foregrounded NARRATIVE clauses — and these findings are parallel across the two generations. The occurrence of *kom* is highly circumscribed for both speakers. Indeed, the correlation of *kom* with sequential narrative complicating action clauses is far stronger than the effects of LEXICAL STATIVITY and AFFIRMATIVE vs. NEGATIVE, neither of which is selected as statistically significant.

Table 2: *Variable rule analysis of the contribution of syntactic/semantic factors to the probability of kom according to generation of the speaker.*

	Adult	Child
CORRECTED MEAN:	.40	.50
TOTAL N:	550	352
FACTORS SELECTED:		
Narrative structure		
Foreground	.52	.55
Background	.22	.14
Range	43	41
Temporal relationship		
Sequential	.67	.62
Non-anterior	.41	.39
Anterior	.14	.17
Range	53	45
FACTORS NOT SELECTED:		
Stative/non-stative	X	X
Disambiguating adverb	X	X
Negation	X	X

Narrative structure

The fact that *kom* is so strongly associated with foregrounded clauses in narrative structure suggests that its function is highly correlated with story-telling. In fact, a number of researchers have already examined patterns of verbal morphology in creole languages in narrative discourse, specifically Bollée (1977), Corne (1977), Rickford (1987b) and LePage & Tabouret-Keller (1985). One of their claims is that the distribution of unmarked verbs in creoles parallels that found for the English Historical Present tense. Givón (1979b), for example, refers to a stylistic/dialectal level in English narrative style which “dispenses with tense”, and where the unmarked form of the verb is used “just like in creoles”. He suggests that the “informal” Standard English rendition of past punctual events in (foregrounded) narrative clauses, which features present-tense marking (i.e. the Historical Present), and the creole zero form (which is predicted to occur in precisely the same clauses), are identical in function (ibid. 128). Moreover, in Givón’s analysis this bi-partite division is thought to reflect the “unmarked case” for discourse-pragmatic organization.

Given these claims for consistently *unmarked* verbs in creole narrative discourse, it is quite surprising to find that the NPE stories in our sample are literally packed with verbs which are *marked* with *kom*. *Kom* occurs extremely frequently, often one after the other in a series of clauses, as can be observed in (8):

- (8) *Wen i kom bɔn fɪnɪʃ, nao,*
i no kom si enitiŋ it
i kom go aut agen
i kom go fɔ dat ples wey i kɔt di tri.
 ‘When it had finished burning, now. He didn’t see anything to eat.
 He went out again. He went to that place where he had cut down the tree.’

However, this is not by any means a categorical pattern. All the complicating action clauses are interspersed with verb stems as well, as can be observed in (9):

- (9) a. *I kom jump wit am.*
Kom jump wit am laik dat
i kom ple.
Na im, wi se “no”.

Wi **kɔm** se “mek wi kari am, they go sey, no wi dɔn tif, o”.

So, wi **kɔm** liv am.

Wi **kɔm** go bak haos,

kɔm go tel awɔ mama, nao.

‘He jumped with us. He jumped with us like that. He played. We said, “no”, we said, “if we take him, they’ll say we stole him”. So, we left him. We went back to the house, told our mama, now.’

- b. Na im a **kɔm** go **haid** di muni nao

na im a tek di muni

go bai sɔmtiŋ

mi and ma sista **kɔm** it di tiŋ. (4/76-7)

‘I went and hid the money now. I took the money, went and bought something. Me and my sister ate the thing.’

In this light, although early work suggested that the patterning between marked and unmarked verb forms could be equated with the French, or English “historical present” (Bollée 1977, Corne 1977), the patterns they described for creoles do *not* in fact correspond precisely to those attested in empirical discourse analyses of the internal structure of narrative (Schiffrin 1981; Silva-Corvalán 1983; Wolfson 1978, 1979, 1981). These studies reveal that it is a specific pattern of *variation* between Historical Present and Past morphology *within* complicating action clauses that characterizes English narratives.

English Historical Present verbs (which are essentially verb stems)⁴ do *not* typically function as the unique marker of foregrounded narrative material (Schiffrin 1981:51). An English narrative usually alternates, in a characteristic distribution pattern, with the morphologically marked past tense (in English the preterit suffix *t, d*, *Id*, or a suppletive vowel change on strong verbs) (see also Tagliamonte & Poplack 1988).⁵ In creoles, on the other hand, there is an apparent contrast between non-marking of sequentially-processed information on one hand, and marking, of whatever type, elsewhere (Givón 1979a: 158).

Looking back at the NPE examples in (9), we observe an alternation between marked and unmarked forms *within* the complicating action of stories reminiscent of that found in English, French, Spanish, etc., rather than the exclusive occurrence of one or the other. Indeed, this tendency toward robust marking of complicating action clauses corroborates Rickford’s (1987a) analysis of Guyanese creole which found such clauses to be marked 40% of

the time. Similar findings are reported in Myhill's (1991) survey and Patrick's (1999) study of Jamaican creole.

Given this discourse-level perspective, the question that arises is what meaning, if any, can be attributed to the *alternation* between *kom*-marked verbs and the stem form in these NPE stories? It is not at all obvious, from a sentence-level perspective alone, exactly what may be underlying this variability.

Thus, the variable context for *kom* was reconfigured in order to examine its function from a discourse level perspective. Following the stipulations in Schiffrin (1982), the analysis is limited to a well-circumscribed discourse context of which *kom* appears to be a particular feature — the genre of story-telling. My intention here is to examine the type of variable patterning that exists between marking and unmarking in this context.

Narrative analysis

The following analyses come from a sub-sample of the data which includes only those verbs found in narratives. This comprised 9 narratives in total, 5 from the child and 4 from the adult. With the exception of one folk tale told by the child, all are stories of personal experience, in the recent life history of the speakers. These materials comprise nearly 900 verb phrases in narrative discourse.

Rather than categorizing verbs according to their temporal sequence and a bipartite division into narrative clauses versus everything else as in Table 2, here each verb was categorized according to its membership in the narrative structure (Labov & Waletzky 1967). First, *complicating action*, or *foregrounded* clauses, as in (10), which comprise the "actual story line" of the narrative and recount what happened in a series of clauses that are temporally ordered according to the sequence in which the events actually occurred:

- (10) a. Ay *kɔm* entɔ
 i *kɔm* put watɔ fɔ mi
 ay bat fɪnɪʃ
 den i *kɔm* mek fud fɔ mi. (13/48-50)
 'I entered. He put water for me. I had a bath. Then he made
 food for me.'
- b. I *kɔm* kɔt am
 kari am go haos
 i *kɔm* plant am fɔ haos. (4/15)
 'He cut them, carried them to the house. I planted them in the
 house.'

Background, or “supportive material” used to “amplify or comment on the events of the main narrative” (Hopper 1979: 213-214) were also distinguished. *Orientation* clauses as in (11), situate the listener with regard to the general circumstances (e.g. time and place) of the events being recounted, as well as the identities of the characters involved. These are similar to stage directions or scene-setting information.

- (11) a. *Enitaim we i deh hɔŋgri,*
i go ʃus go deh
go tek banana
stɔt deh it im. (4/17)
 ‘Anytime that he was hungry, he would just go there and go take a banana and start to eat it.’
- b. *Wi dɔn deh ple awɔ mjuzIk.*
wi deh ʃat.
wi deh laf.
ay no evIn flt luk arawn bekɔs na naitaim. (13/45-6)
 ‘We had been playing our music. We were chatting. We were laughing. I didn’t even look around because it was night time.’

Evaluation clauses communicate to the listener the narrator’s feelings about the events (i.e. provide the “point” [Polanyi 1979] of the narrative), as in (12). These are contexts in which the narrator steps out of the story and says how she/he feels or interprets events:

- (12) a. *Wi kɔm hapi*
sey wi dɔn even rič
wea wi deh go at las. (13/36)
 ‘We were happy that we had found where we were going at last.’
- b. *I no no wetin i go do.* (4/12)
 ‘He didn’t know what he was going to do.’
- c. *Im own luk kɔm ʃain fɔ dat de.* (4/269)
 ‘His luck shone that day.’

Finally, we distinguish *abstract* and *coda*, which are the metalinguistic brackets which introduce a narrative, as in (13), and bring it to a close, as in (14):

- (13) a. ... *dat time wey my mama leave me go Canada*
dat time a kɔm deh cry
until wen mi mama kɔm
we kɔm deh hapi
evribodi kɔm deh hapi. (4/186-191)
 ‘That time, when my mom left me to go to Canada. That time, I
 cried until my mom came back. Then we were happy. Every-
 body was happy.’
- b. *Dat wun na big wahala, oh!* (N/111)
 ‘That one was a big problem, oh!’
- (14) a. *Dat tory no sweet you for ear?* (N/195)
 ‘Wasn’t that a good story?’
- b. *Na deh da story kɔm flɪnlɔ̃.* (4/121)
 ‘This is where the story ends.’

Table 3 shows the distribution of verb forms according to narrative structure:

Table 3. Distribution of verb forms in each narrative section

Verb form:	Complicating Action	Orientation	Evaluation	Abstract/ Coda	Total N
	% N	% N	% N	% N	N
<i>kɔm</i> + verb	58 213	12 53	5 3	7 1	270
<i>kɔm</i> + <i>deh</i> + verb	5 18	5 23	—	13 2	43
verb stem	30 110	38 163	29 17	27 4	294
verb stem (elision context)	8 28	2 7	—	—	35
<i>deh</i> + verb	—	12 52	14 8	13 2	62
copulas	—	9 39	29 17	33 5	61
<i>dɔn</i> + verb	—	4 16	7 4	—	20
<i>go</i> + verb	—	18 77	16 9	7 1	87
<i>mek</i> + verb	—	1 4	—	—	4
TOTAL N	369	434	58	15	872

The *orientation* clauses and *evaluation* clauses contain the widest variety of verb forms, including nearly all of the *don* and *deh*-marked verbs. The *complicating actions* are marked with *kom* nearly two-thirds of the time (231/369). Indeed, nothing else occurs in the *complicating action* except the verb stem. Moreover, narrative clause *kom* constitutes a full 79% (213/270) of all the *kom* + verb constructions (213/270). What this partitioning demonstrates is that there is a tremendous difference in verb morphology between iconically-ordered foregrounded narrative clauses on one hand, versus background clauses (i.e. *orientation* and *evaluation*), which contain all the special markings required to interpret out-of-sequence, non-punctual and irrealis states and events.

Thus *kom* is a highly circumscribed form, occurring overwhelmingly in a well-defined portion of narrative structure, rather than randomly throughout — the *complicating action*. Crucial, however, is the fact that this narrative component is not at all uniquely marked with *kom*, nor unmarked, as early work on pidgins and creoles have found. Instead, *kom* and the verb stem alternate, just as found by Rickford (1987a), Myhill (1991), and Patrick (1999).

This variability between zero and *marked* verbs in complicating action clauses exactly parallels the type of surface alternation which, as we have seen, has been reported for many other languages by Schiffrin 1981, Wolfson, 1981, Silva-Corvalán 1983, Tagliamonte & Poplack 1988. I now examine the distribution of verbs marked with *kom* and those surfacing as verb stems and compare them with *patterns* that have been reported for tense alternation phenomena in the literature on this subject. Since Schiffrin (1981) provides the most replicable analysis, I use this as a point of cross-linguistic comparison.

First, we will consider the *rate* of verb marking in complicating action clauses compared to other narrative sections. The *kom*-marked verbs are concentrated almost exclusively in the complicating action clauses of the narrative, illustrated in Table 4, where they represent approximately two thirds of all verbs. Moreover, the proportions are nearly identical between the two speakers. The rate is much attenuated in all other narrative sections. While this is consistent with the findings in Table 1, notice that the *rate* of marking corresponds very closely to the typical marking rate of verbs in English story-telling mode:

Table 4. Distribution of “marked” verbs in narrative sections

Data:	Complicating Action		Orientation		Evaluation		Abstract/Coda	
NPE <i>kom</i>	%	N	%	N	%	N	%	N
Adult	69	174	19	207	4	52	—	9
Child	57	195	15	227	17	6	50	6
English <i>-ed</i> (Schiffrin 1981)	70	1288						

Next, illustrated in Table 5, note the tendency to begin and end the complicating action section with marked (past tense) verbs, while the Historical Present occurs most often in middle clauses:

Table 5. Location of verb stems in complicating action clauses

	Initial Clause		Middle Clauses		Final Clause	
NPE verb stem	%	N	%	N	%	N
Adult	0	4	32	167	50	4
Child	0	5	46	184	40	5
English verb stem (Schiffrin 1981)	18		32		9	

Unfortunately, the sparse data for initial and final clauses for the complicating action of the narratives makes it impossible to arrive at any conclusive statement on the patterning here. It is merely suggestion, therefore, that in initial complicating action clauses the verbs are unanimously marked with *kom*, paralleling the strong tendency for marking in this context in Schiffrin’s study. More relevant is that rates of marked verbs in middle clauses hover around the same frequency across studies.

Another characteristic pattern of the Historical Present/Past tense alternation is that verbs of the same tense tend to cluster together, making rapid alternation between the verb stem and the *kom*-marked verb atypical (Schiffrin 1981:51). Table 6 demonstrates this, in showing that both verb stems and *kom*-marked verbs are more frequent when the prior verb is marked with *kom*:

Table 6. Clustering of verb forms in complicating action clauses

		VERB STEM AFTER VERB STEM		MARKED VERB AFTER MARKED VERB	
NPE	<i>kom</i>	%	N	%	N
	Adult	42	50	69	101
	Child	54	79	60	103
English -ed					
(Schiffrin 1981)		62	368	82	847

Although this clustering effect is more pronounced in the adult than the child, both NPE speakers exhibit the same direction of effect.

Taken together these results reveal striking parallels between marked and unmarked forms in the NPE stories and previous quantitative studies of creoles. Moreover, the *patterns* of alternation are similar to those found for the Historical Present and simple past tense found in English stories. It does not follow, however, that NPE is behaving like English. Patterned alternation between overt verbal morphology of this type is also characteristic of French and Spanish stories as well. More to the point, qualitatively, the forms which represent the alternation in each of these languages are entirely different. For example, in French the morphological alternation is between the suffixes of the simple present tense and the auxiliary and verbal endings of the *passé composé*. On the contrary, NPE's marked form, pre-verbal *kom*, is unlike any verbal marker in any Indo-European language, but a common lexical source of past tense markers in many indigenous Nigerian languages (Heine & Reh 1984). Thus while the patterns are the same across varieties, the *forms* are different. However, as Fleischman (1985:8) suggested with regard to this so-called tense switching phenomenon, there is no reason to expect universally consistent mapping onto the same grammatical categories despite the fact that the discourse-pragmatic mechanism might remain the same.

Leaving aside the question of the source of these forms and their patterns, the systematic nature of the alternation between *kom* and the verb stem revealed in this section provides evidence for the fact that at the level of discourse as well, NPE *kom* exhibits qualities of being a highly grammaticalized feature of the language. It encodes not only syntactic/semantic meaning, but also discourse-level meaning. Here too the patterning between adult and child speakers is nearly identical on each point of comparison.

Information structure

In addition to narrative structure, there is another aspect of discourse organization which has been found to bear considerably on the interpretation and referential function of linguistic features. Schifffrin (1987) has demonstrated that language exhibits a sophisticated degree of structure over large stretches of discourse in which the coherence and hierarchical arrangement of information in the progression of the discourse may contribute to the encoding of referential function. This organization can be observed in the strategic placement and co-occurrence patterns of what she refers to as “discourse markers”. These markers work to subordinate and coordinate various levels of discourse as well as to express interactional alignments and convey social meanings. Discourse markers in particular can serve as “rudders” to guide the processing of language in order to give cohesion to levels of interpretation of discourse and signal hierarchies or chunks of information (Ferrara 1997). The fact that discourse markers have this double-duty both as organizational units and as features expressing referential meanings suggests that their distributional patterns may reveal the ways in which discourse level features interact with grammatical elements in the synchronic organization of language.

In fact, stories of personal experience typically contain many digressions from the main storyline as narrators provide further information, give scene-setting details, look back to earlier times, and get interrupted. For example, Ferrara (1997) observes that the form *anyway* connects two levels of representation, and that resumptions could span large passages of intervening text in personal narratives, whereby the marker demonstrates cohesion not with the preceding sentence, but with material many lines away. Such empirical findings suggest that more than just verb forms are involved in marking the distinction between foreground and background. As one of the most salient functions of discourse markers is to subtly signal a resumption of the trend of thought of the speaker or narrator (ibid. 9), each context was coded for its position in the informational structure of the narrative.

Table 7 considers what nuances of verbal patterning may be revealed through an examination of this macro-level view of the stories:

Table 7. Location of *kom*-marked verbs and verb stems by information structure of the narrative

	1ST COMPLIC. ACTION CLAUSE	1ST CLAUSE AFTER BACKGR' ND	INTERNAL CLAUSE	FINAL COMPLIC. ACTION CLAUSE	BACKGR' ND CLAUSES
	% N	% N	% N	% N	% N
Child					
<i>kom</i> + verb	100 5	88 36	46 62	60 3	33 45
Adult					
<i>kom</i> + verb	100 4	88 43	61 68	50 2	29 45
TOTAL N	9	90	247	9	290

The table shows that *kom* tends to occur at very precise points in the narrative structure. It appears a full 88% of the time at points in the stories which encode resumptions to the main storyline. Internal narrative clauses are much less likely to be marked with *kom* and backgrounded clauses even less. Thus, *kom* is highly correlated with information management in the narratives. Again the parallels between adult and child in each of these discourse environments are striking.⁶

While research on discourse-level phenomena has primarily focused on the form and function of the discourse markers themselves, here I take an alternative perspective and examine their specific placement and co-occurrence patterns along with verbal morphology. Although such collocations are not necessarily indicative of meaning, as Schifffrin (1987) points out, they can serve as *pointers* to the functions of grammatical features.

Accordingly, each clause was coded for markers that might be implicated in the organization and/or referential meaning of the discourse, such as *so*, as in (15), *na im* 'so' or 'then', an idiosyncratic conventionalization often used in narratives (Charles Mann, p.c., January 1999), as in (16), *now*, as in (17), as well as adverbs, as in (18) and temporal conjunctions, as in (19):

- (15) a. *So, de kom dlrek mi.*
 'So, they directed me.'
 b. *Na so wun man kom.* (4/100)
 'So, one man came.'

- c. *So, wen wi kɔm entɔ fɔ insaid de nao.* (13/10)
 ‘So, when we entered inside there now.’
- (16) a. *Na im e kɔm waka go hom.*
E waka flɪŋʃ.
Na im e no see enitin. (4/11-2)
 ‘He walked home. His walk finished. He didn’t see anything.’
- b. *Na im a ʃus tek ma pikčɔ deh go.* (13/114)
 ‘I just took my picture and went.’
- c. *Na im wi kɔm go slip nao.* (13/59)
 ‘We went to sleep now.’
- (17) a. *Wen krɪsmɔs de kɔm rič nao, e kɔm deh sɪk.* (13/244)
 ‘When Christmas day arrived now, he was sick.’
- b. *Wi kɔm begin entɔ sɔm gas staʃɔnz, kɔm deh aks sɔm pipl nao.*
 (13/28-9)
 ‘We began entering some gas stations, asking people now.’
- (18) a. *Di neks de wi kɔm dewaka araown.* (13/70)
 ‘The next day we were walking around.’
- b. *Dat taim wey ma mama liv mi go kanada. Dat taim a kɔm deh krai* (4/186)
 ‘The time that my mother left me to go to Canada. That time I was crying.’
- c. *Den wi kɔm begɪn nok tori o.* (13/57)
 ‘Then we began to tell stories.’
- (19) a. *Wen ma mama kɔm tel ɔs se im get aksɪden fɔ rod...* (13/192)
 ‘When my mother told us that she had an accident on the road ...’
- b. *As a kɔm araiv nao.* (13/6)
 ‘As I arrived now.’

The sheer frequency, and co-occurrence of these markers in the narratives, as in (20), is remarkable:

- (20) *[na im] [wun dai] [nao] ma brodɔ kɔm gɪv mi muni.* (4/73)
 ‘One day my brother gave me money.’

In order to determine which, if any, of these discourse-level correlations are significant to the appearance of *kom* when all the factors are considered simultaneously, we turn again to multivariate analysis.

5. Multivariate Analysis

Table 8 reconfigures the data shown earlier in Table 2 according to narrative and information structure and the other discourse/pragmatic factors.⁷ The results depict the probability of *kom* in the child and adult speaker when all these factors are considered simultaneously.

Table 8. Variable rule analysis of the contribution of discourse factors to the probability of *kom* in Nigerian Pidgin English according to generation of the speaker

	Adult	Child
CORRECTED MEAN:	.324	.313
TOTAL N:	453	444
FACTORS CONSIDERED:		
Information structure		
First in foreground	100% (N=4)	100% (N=5)
Foreground clause after background	.93	.93
Internal clause	.73	.60
Background clause	.30	.34
Last in complicating action	50% (N=4)	60% (N=5)
Range	63	59
Mark on preceding reference verb		
Preceding <i>kom</i>	.58	.60
Preceding verb stem	.47	.49
Other preceding mark	.38	.34
Range	20	26
Phrase initial <i>so</i>		
<i>so</i>	.79	KNOCK-OUT
no introducer	.47	100% (N=3)
	32	
Phrase initial <i>na im</i>		
<i>na im</i>	[.64]	.75
no introducer	[.50]	.48
		27
FACTORS NOT SELECTED:		
Disambiguating adverb	X	X
Temporal conjunction	X	X
Co-occurring <i>now</i>	X	X
Phrase initial <i>na im</i>	X	
Phrase initial <i>so</i>		n/a

For both speakers, the INFORMATION STRUCTURE of the discourse is the greatest and most significant discourse-level effect conditioning *kom*. There is also a statistically significant effect of MARK ON THE PRECEDING REFERENCE VERB, with a clear tendency toward clustering of the overtly marked *kom* verbs. The direction of effect of both the INFORMATION STRUCTURE and CLUSTERING factors as well as their relative strength vis-à-vis the other factors is parallel between the adult and child.⁸ Local adverbial and/or temporal conjunctions exert no statistically significant effect on *kom*. These findings reveal that *kom*-marked verbs are strongly implicated in the macro-level organization of NPE discourse. Moreover, this is once again parallel across these two generations of speakers.

Notice, however, the very different patterns between child and adult with regard to collocation patterns with the two discourse markers *na im* and *so*. What is relevant here is not that the constraint hierarchy is similar, since the data are very sparse. Instead it is notable that the child uses only three tokens of *so*, while the adult uses far more (N=35). The differential effect of these two discourse markers is the first result which clearly distinguishes the adult and child. Why would this be so? It is undoubtedly not coincidental that *so* is a discourse marker in English. Although there are no available analyses which could directly correspond with those conducted here, in previous empirical work it has been identified as a marker of “main idea units” (Schiffrin 1987:91). Thus, instead of a deep-rooted grammatical difference between these two speakers, the preponderance of *so* and *na im* as well as their very similar distribution pattern suggests a functional slot which is parallel across NPE and English. The adult’s use of *so* may simply be an importation of a discourse marker from English, given her exposure to the local vernacular in Ottawa over a year. This interpretation is supported by recent research by Sankoff et al. (1997) which demonstrates that fluency in a local vernacular entails the use of the community’s repertoire of discourse markers. In the same article Sankoff et al. (1997) note the extensive use of *so* among French minority speakers in Ontario (Mougeon & Beniak 1991) and in New Brunswick (Roy 1981). Use of *so* is also reported with high frequency in French minority speakers in Nova Scotia (Flikeid 1996). The consistency of these reports suggests that *so* is a discourse marker which is easily transferred from Canadian varieties of English. The fact that this speaker uses both *na im* and *so*, indeed in some contexts using both of them, as in (21), suggests the parallelism of the two forms in her repertoire:

- (21) a. [so] [na im] i kari di got ... (13/187)
 ‘So, he took the goats...’
 b. [so] [nao] [na im] wi kom kari awɔ got, go (13/195)
 ‘So now, we took our goat and went.’

Given these findings, it might be hypothesized that if functional parallelism obtains at the discourse level, this may be a prime location for the importation of linguistic material from one language to another. More generally this may be an indication of the conditions of mutual reinforcement that initially favor transfer in language contact situations. Parallelism between functions provides convergence which favors the change of *na im* and *so* away from their literal meanings into more grammatical interpretations. However, further empirical and comparative work on discourse markers in language contact situations, particularly in the function of *so* in both English and NPE as well as *na im*, is needed to determine which of them will gain selective advantage over the other alternatives, consistent with Mufwene’s (1996b) competition-and-selection approach which operates under the specific constraints of local ecological conditions.

6. Summary

In sum, all the findings reported here suggest that *kom* is deeply entrenched in all levels of NPE grammar: syntactic, semantic and discursive. Thus, returning to the question of whether *kom* is “rapidly grammaticalizing” in NPE, it is perhaps now more appropriate to ask whether this form, which seems to have grammaticalized already, could still reveal the tracks of its path of change. Alternatively, has the grammatical category simply been transferred into NPE from English or its substrate languages?

Taking Faraclas’ (1987: 55, Table 1) tabulations as an earlier representation of the frequency of *kom* at approximately 30% in L1 speakers c. 1985, it appears that *kom* has increased substantially in frequency in the last decade, which suggests an explanation for its more advanced grammatical status in the current data set. As is well-known in the literature, the changes associated with grammaticalization are actually a continuum in which different degrees of restructuring may exist at different points in time as the original lexical content of the form gradually bleaches. This is what Hopper (1991:23) has referred to as “layering”. Moreover, even a highly grammaticalized form

might still show evidence of incomplete desemanticization if it is the product of an extended process in diachronic time since the original meanings of forms “tend to persist over time and to constrain the later uses of the grammaticalized form” (Hopper & Traugott 1993:93). If so, what might be the last bastion of influence of the original lexical content of the form *kom*?

Perhaps the best evidence of nearly complete grammaticalization of *kom* is the complete loss of its original lexical meaning. This could be observed by demonstrating that *kom* is as can co-occur with the lexical verb *kom* without creating redundancy, as shown in (22) (and earlier in [6]):

- (22) a. *Na im pusi kom kom aut,*
tel am, se “oh oh, yu no go si eniwɔn it agen.” (4/034)
 ‘Pussy, came out and told him, “oh, oh, you aren’t going to see anyone eat again.’
- b. *Aftɔ de papa and de mama kom kom bak frɔm wɔk ...* (4/140)
 ‘After their papa and their mama came back from work ...’

Table 9 presents the distribution of *kom* by the lexical verb *kom* vs. other verbs:

Table 9. Distribution of *kom* by main verb types across two generations

Main verb	CHILD		ADULT	
	%	N	%	N
<i>kom</i>	36	28	17	12
other	34	358	36	359

Here, despite the small Ns, observe that there is an almost two-fold increase in the child’s use of auxiliary *kom* with the main verb *kom* compared to that of the adult, as well as that the choice of auxiliary *kom* is as likely with main verb *kom* as any other. Thus, in this one very limited area we find some evidence which is suggestive of *kom*’s grammaticalization pathway: here we find the child does indeed represent a somewhat farther point along the grammaticalization cline than the adult.

7. Conclusions

In nearly every analysis to which it has been subjected *kom* exhibits remarkably similar patterns and constraints across speakers which may be taken to

represent two recent generations of NPE. These findings corroborate and bolster the findings of earlier work in demonstrating that *kom* is a central part of NPE's relative tense system serving to denote sequential past time.

Detailed comparison of distribution patterns demonstrates that alternation between *kom* and the verb stem in narrative discourse is the norm in NPE. This differs from the creole patterns reported in early work but provides strong corroborating support for the marked and unmarked verb patterns in narrative discourse found in other quantitative studies of creole languages. The results also demonstrate that this alternation parallels the patterning found in discourse analyses of narratives in other languages. Thus, while NPE *kom* is clearly a marker of SEQUENCED temporal relationships in a manner distinct from that of its lexifier, it participates in macro-level discourse organization that is quite similar to the HISTORICAL PRESENT/SIMPLE PAST TENSE alternation in Indo-European languages. Whether this is the result of universals of discourse organization, or a reanalysis of the same phenomenon in its lexifier language must await comparative empirical studies of such discourse phenomena in non-Indo European languages.

The sum total of these analyses, examining factors from different areas of the grammar, demonstrates that *kom* is a highly grammaticalized form, exhibiting complex multi-level variability conditioned by features of semantics, syntax and discourse. Moreover, with the exception of the adult speaker's use of the English discourse marker *so*, these speakers show very little influence, if any, from English, despite the fact that the adult is a highly educated speaker of NPE currently living in a predominantly English-speaking community.

As to how discourse effects are implicated in the process of change, these findings provide little evidence, since *kom* is unfortunately beyond the level at which such effects might be observed. The discourse markers found in these data serve more as reinforcement and/or discourse management devices to aid in the navigation of the discourse. However, the fact that they are so strongly associated with the distribution patterns of *kom* suggests that they may have been implicated in earlier stages of its development. Comparison of the distribution of such patterns at earlier points in the trajectory of a grammaticalizing form would thus be a fruitful area of research to pursue in examining the contributions of discourse to grammatical change.

Finally, these findings point to the fact that grammaticalization, even at an advanced stage, can still be observed in quantitative patterns of more or less. Even beyond the point at which variable constraints at the syntactic and

discourse components of the grammar may be differentiated across generations, the pathway of change in the history of a grammaticalizing feature may still be apparent at the lexical level. If these findings are any indication, however, the story of *kom* in Nigerian Pidgin English, at least with regard to its state of grammaticalization, is by now pretty much complete.

Notes

- * This research was made possible through the generous support of the Social Sciences and Humanities Research Council of Canada in the form of research grant #410-90-0336. I thank Ejike Eze for his assistance in collecting and transcribing the data and for helping me with the English translations. I am grateful to R. B. LePage, Salikoko Mufwene, and an anonymous reviewer for their insightful comments on an earlier version of this manuscript.
- 1. Indeed, *kom* occurs much more frequently in past temporal reference than *bin*, a finding that may seem unusual given the literature on English-based creoles. For further discussion and a detailed analysis of the NPE past temporal reference system see Poplack & Tagliamonte (1996b).
- 2. Interestingly, in Igbo, the first language of these speakers, verbs surface bare in negative sentences with past temporal reference (see discussion Poplack & Tagliamonte 1996b).
- 3. These two factors are partially overlapping. Narrative complicating action clauses are, with the exception of restricted narrative clauses, sequentially ordered (Labov & Waletzky 1967). However, backgrounded clauses may embody any temporal relationship.
- 4. Because English has an overt present tense morpheme only in the third person singular, the Historical Present is for all intents and purposes a “stem” form because the verbs surface without overt morphology.
- 5. While it is possible that English narratives may be told entirely in one tense (Schiffrin 1981, Le Page & Tabouret-Keller 1985), I emphasize here the consistency of patterned variation between historical present and past morphology revealed in empirical quantitative analyses in the literature.
- 6. The child’s heightened rate of verb stems in internal narrative clauses is due to the preponderance of contexts, as in:
 - (i) *e ĵus kom take di rabbit frɔm am kɪl di monki put am fɔ insaid bag cari am go.*
 ‘He just took the rabbit from him, killed the monkey, put it in the bag, carried it, went.’ (4/45)
- 7. The difference between Table 2 and Table 8 is that Table 8 tests a detailed categorization schema of narrative discourse structure.
- 8. Thus, the earlier difference in clustering patterns across the two speakers does not turn out to be statistically significant.

References

- Aghegyisi, R. N. 1984. "Linguistic implications of the changing role of Nigerian Pidgin English". *English World-Wide* 52: 211-233.
- Baker, P. and Syea, A. (eds) 1996. *Changing meanings, changing functions. Papers relating to grammaticalization in contact languages*. Westminister, UK: University of Westminister Press.
- Bollée, A. 1977. *Le créole français des Seychelles*. Tübingen: Gunter Narr.
- Bruyn, A. 1995. *Grammaticalization in creoles: The development of determiners and relative clauses in Sranan*. Amsterdam: IFOTT.
- Corne, C. 1973. "Tense and aspect in Mauritian creole". *Te Reo* 16: 45-59.
- Corne, C. 1977. *Seychelles Creole grammar*. Tübingen: Niemeyer Verlag.
- Faraclas, N. G. 1987. "Creolization and the tense-aspect-modality system of Nigerian Pidgin". *Journal of African Languages and Linguistics* 91:45-59.
- Ferrara, K. 1997. "Form and function of the discourse marker *anyway*: Implications for discourse analysis". *Linguistics* 35: 343-78.
- Fleischman, S. 1985. "Discourse functions of tense-aspect oppositions in narrative: Toward a theory of grounding". *Linguistics* 23: 851-882.
- Flikeid, K. 1996. *Code-switching as a pathway to structural borrowing? A French-English multi-community study*. Paper presented at Sociolinguistics Symposium 11, University of Wales, Cardiff, UK.
- Givón, T. 1979a. "Tense-aspect-modality: The creole prototype and beyond". In *Tense-aspect: Between semantics and pragmatics*, P. Hopper (ed.), 115-63. Amsterdam: John Benjamins.
- Givón, T. 1979b. "Prolegomena to any sane creology". In *Readings in creole studies*, I. Hancock (ed.), 3-35. Ghent: Story-Scientia.
- Givón, T. 1984. "The pragmatics of referentiality". In *Georgetown University Round Table on Language and Linguistics 1984. GURT '84. Meaning, form, and use in context: Linguistic applications*, D. Schiffrin (ed.), 120-38. Washington, D.C.: Georgetown University Press.
- Heine, B. and Reh, M. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.
- Hopper, P. 1991. "On some principles of grammaticization". In *Approaches to Grammaticalization*, E. Traugott and B. Heine (eds), 17-35. Amsterdam: John Benjamins.
- Hopper, P. and Traugott, E. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Hopper, P. J. 1979. "Aspect and foregrounding in discourse". In *Syntax and semantics: Discourse and syntax*, T. Givón (ed.), 213-41. New York: Academic Press.
- Labov, W. and Waletzky, J. 1967. "Narrative analysis: Oral versions of personal experience". In *Essays on the verbal and visual arts*, J. Helm (ed.), 12-44. Seattle: University of Washington Press. 12-44.
- Le Page, R. B. and Tabouret-Keller, A. 1985. *Acts of identity: Creole based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Mougeon, R. and Beniak, É. 1991. *Linguistic consequences of language contact and restriction: The case of French in Ontario, Canada*. Oxford: Oxford University Press.

- Mufwene, S. S. 1996a. "Creolization and grammaticization: What creolistics could contribute to research on grammaticization". In Baker and Syea, 5-28.
- Mufwene, S. S. 1996b. "The founder principle in creole genesis". *Diachronica* 13: 83-134.
- Myhill, J. 1991. "Typological text analysis: Tense and aspect in creoles and second languages". In *Crosscurrents in second language acquisition and linguistic theories*, T. Heubner and C. A. Ferguson (eds). 93-121. Amsterdam: John Benjamins.
- Patrick, P. 1999. *Urban Jamaican Creole: Variation in the Mesolect*. Amsterdam: John Benjamins.
- Plag, I. 1998. "On the role of grammaticalization in creolization". Paper presented to the Society for Pidgin and Creole Linguistics, Symposium on Pidgin and creole linguistics in the 21st century, New York.
- Polanyi, L. 1979. "So what's the point?" *Semiotica* 25: 207-241.
- Poplack, S. and Tagliamonte, S. 1995. "It's black and white: The future of English in rural Nova Scotia". Paper presented at the NWAVE 24, University of Pennsylvania, Philadelphia.
- Poplack, S. and Tagliamonte, S. 1996a. "The grammaticization of *gonna* in six varieties of English: A cross-linguistic comparison". Paper presented at NWAVE 25, University of Nevada, Las Vegas.
- Poplack, S. and Tagliamonte, S. 1996b. "Nothing in context: Variation, grammaticization and past time marking in Nigerian Pidgin English". In Baker, 71-94.
- Rand, D. and Sankoff, D. 1990. *GoldVarb. A variable rule application for the Macintosh*. Montreal, Canada: Centre de recherches mathématiques, Université de Montréal. Version 2.
- Rickford, J. 1987a. *Dimensions of a creole continuum: History, texts, and linguistic analysis of Guyanese Creole*. Stanford, California: Stanford University Press.
- Rickford, J. 1987b. "Past marking in the Guyanese mesolect: A close look at Bonnette". In *Variation in language*, K. Denning, S. Inkelas, F. McNair-Knox and J. Rickford (eds.), 379-94. Stanford, California: Stanford University.
- Roy, M.-M. 1981. "Les conjonctions BUT et SO dans le français de Moncton". In *Variation omnibus*, D. Sankoff and H. Cedergren (eds), 429-44. Edmonton, Alberta: Linguistic Research.
- Sankoff, G. 1990. "The grammaticalization of tense and aspect in Tok Pisin and Sranan". *Language Variation and Change* 2: 295-312.
- Sankoff, G. and Brown, P. 1976. "The origins of syntax in discourse". *Language* 52: 631-666.
- Sankoff, G., Thibault, P., Nagy, N., Blondeau, H., Fonollosa, M.-O. and Gagnon, L. 1997. "Variation in the use of discourse markers in a language contact situation". *Language Variation and Change* 9: 191-217.
- Schiffrrin, D. 1981. "Tense variation in narrative". *Language* 57: 45-62.
- Schiffrrin, D. 1982. *Discourse markers*. University of Pennsylvania Ph.D. dissertation.
- Schiffrrin, D. 1987. *Discourse markers*. Cambridge: University of Cambridge Press.
- Silva-Corvalán, C. 1983. "Tense and aspect in oral Spanish narrative: Context and meaning". *Language* 59: 761-80.
- Singler, J. V. 1992. "Nativization and pidgin/creole genesis: A reply to Bickerton". *Journal of Pidgin and Creole Linguistics* 7: 319-333.

- Tagliamonte, S. and Poplack, S. 1988. "How Black English *past* got to the present: Evidence from Samaná". *Language in Society* 17: 513-33.
- Tagliamonte, S., Poplack, S. and Eze, E. 1997. "Pluralization patterns in Nigerian Pidgin English". *Journal of Pidgin and Creole Languages* 12: 103-29.
- Traugott, E. C. 1982. "From propositional to textual to expressive meanings: Some semantic-pragmatic aspects of grammaticalization". In *Perspectives in historical linguistics*, W. P. Lehmann and Y. Malkiel (eds.), 245-71. Amsterdam: John Benjamins.
- Traugott, E. C. and Heine, B. 1991. *Approaches to Grammaticalization*. Amsterdam: John Benjamins.
- Wolfson, N. 1978. "A feature of performed narrative: The conversational historical present". *Language in Society* 7: 215-237.
- Wolfson, N. 1979. "The conversational historical present alternation". *Language* 55: 168-182.
- Wolfson, N. 1981. "Tense-switching in narrative". *Language and Style* 14: 226-31.

Tense and Aspect in Sranan and the Creole Prototype

Donald Winford

1. Introduction*

My primary aim in this paper is to reconsider the notion of a “prototypical creole TMA system” as defined by Bickerton (1974, 1981) in the light of a fairly comprehensive analysis of the tense/aspect system of Sranan (henceforth SN), based on a large database of recorded speech. The analysis is intended to act both as a corrective to previous descriptions of this creole and as a model for descriptions of creole TMA systems in general. It is hoped that this alternative approach will allow us to address the question of creole similarities and differences more precisely, as a prerequisite for addressing the issue of creole origins, as well as the relative roles of universals, superstrate inputs and substrate influence in creole formation.

For over two decades now, the conventional wisdom in the field has accepted Bickerton’s notion of a “prototypical” creole TMA system. According to him, the TMA system of a “radical” creole comprises two components: an inventory of three categories (an anterior tense, an irrealis mood and a non-punctual aspect) and an invariant ordering of tense, mood and aspect. Most descriptions of creole verb complexes have focussed primarily on to what extent they match this prototype. Such approaches, while they provide valuable information, are limited because the prototype itself is questionable, and preoccupation with it distracts from the task of describing each creole TMA system faithfully and on its own terms. The present study demonstrates that even Sranan, one of the putative exemplars of the prototype, does not conform to it in all respects. The discussion here will be confined to the tense and aspect categories of Sranan.

In fact, Bickerton himself used Sranan to illustrate his prototypical TMA

system, as shown in Table 1. Both the prototype itself and the bioprogram hypothesis associated with it have come under increasing challenge of late, and it has become clear that neither can continue to be accepted in their traditional form. Singler (1986, 1990) provides a lucid summary of the controversy created by the LBH and the various objections that have been made against it on both sociohistorical and linguistic grounds, so I will not attempt to reproduce that discussion here. Various other researchers, including Boretzky (1983), Muysken (1981), Spears (1990), Gibson (1992), Jaganauth (1987), Winford (1993) have also pointed to the ways in which many creole TMA systems in fact depart from Bickerton's putative prototype.

Table 1. Bickerton's 'prototypical' creole TMA system (illustrated with Sranan markers)

	Anterior	Irrealis	Non-punctual	
----->	- ----->	- ----->	- ---->	V
			+ ---->	e V
	----->	+ ----->	- ---->	sa V
			+ ---->	sa e V
----->	+ ----->	- ----->	- ---->	ben V
			+ ---->	ben e V
	----->	+ ----->	- ---->	ben sa V
			+ ---->	ben sa e V

One of the recurring themes of the criticism levelled against Bickerton's prototype is the need to question the very accuracy of his analysis of the core categories themselves. In the first place, as I shall argue, even the designations Bickerton employs ('anterior, irrealis and nonpunctual') are questionable, both from the point of view of usual terminological practice in the field, and with respect to the putative meanings and functions associated with the categories themselves. I will argue that Bickerton's prototype is in fact based on questionable analyses of the TMA systems of the paradigm cases. To some extent, this may be explained by the fact that until recently, none of these systems had been thoroughly investigated, and only their general outlines were known. This is true of creole TMA systems in general, few of which have been comprehensively described. Hence debate on issues such as the creole prototype and the origins of creole TMA systems has proceeded on the basis of partial descriptions of a few creoles. For instance, of the 20-odd English-lexicon Caribbean creoles, only Guyanese Creole (henceforth GC) and to a lesser extent SN have figured prominently in discussion of these

issues. Moreover, Bickerton's understanding of the TMA systems of these creoles was based on inadequate accounts which have since been superseded by more detailed analyses. Taking GC as an example, recent work by Gibson (1986, 1992), Jaganauth (1987, 1988), Rickford (1987) and Winford (1993) have provided a much clearer picture of its verb complex than is to be found in Bickerton (1975).

These more recent studies confirm much of Bickerton's analysis (on which he partly based his conception of the creole prototype). There is agreement, for instance, that auxiliaries *bin*, *go* and *a* form the core of the system, though not all researchers accept Bickerton's designations for these categories. In addition, there is agreement with much of Bickerton's analysis of the basic meanings and uses of the categories. At the same time, however, several other aspects of his analysis have been challenged, as will be discussed within the context of further challenges to be outlined in this paper.

Despite all the shortcomings that have been pointed out in it, Bickerton's analysis of the 'prototypical' creole TMA system, and particularly his account of the GC verb complex (1975), were important and highly insightful first steps toward specifying both the similarities and differences among creoles in this area of grammar. Though his explanation for the similarities in terms of the LBH are no longer generally accepted, it still raised crucial questions about creole genesis and the relationships among creoles that are still being debated today. It is not my intention here to discuss alternative explanations for the well-known similarities and differences among creoles. My main aim in this paper is to provide a basis for accurate and comprehensive analyses of contemporary creole TMA systems which will allow us to make accurate comparisons across creoles and ultimately draw inferences about their relationships and origins.

The need for such an approach is made more pressing by the fact that available analyses of creole TMA systems employ different terminology, theoretical assumptions and methods of data collection, and this makes the task of comparison across creoles all the more difficult. Many disagreements about analysis of particular categories seem to be due to lack of consensus on terminology. For instance, the categories referred to as "anterior" and "non-punctual" by Bickerton have also been labelled "relative past" and "imperfective" respectively by other researchers. B's "completive" is also referred to as "perfective-resultative" (Stolz 1987), and so on. There are also disagreements and inconsistencies, as we have seen, in describing the "core meaning" of

certain categories such as “anterior” and “irrealis”. These examples only highlight what is in fact a very serious problem of terminological confusion in the creole literature. Part of the problem, it seems, is due to the fact that creolists employ quite different labels from those used in the more general literature on TMA systems.¹ This paper is partly an attempt to bring approaches to this area of creole grammar in line with those generally employed in the field of studies of tense, mood and aspect (henceforth TMA).

The framework used in the present study employs a set of labels and definitions which are widely accepted in the typological literature on the semantics of TMA systems cross-linguistically. It is closely modelled upon Dahl’s (1985) and Bybee et al’s (1994) studies of a wide variety of languages, a primary aim of which was to seek cross-linguistic generalizations about TMA systems. Dahl’s work in particular provided a model for the present research. His (1985) study investigates tense and aspect categories in a sample of 64 languages representing 11 ‘areal-genetic’ groups. His chief objective was ‘to determine in what ways [TMA] systems in human languages are similar to each other, and in what ways they may differ’ (1985:31). Apart from providing a coherent framework for investigating tense/aspect categories and their meanings, Dahl also provided an exhaustive questionnaire used for elicitation of data, a modified version of which was employed for the present study.

2. A framework for analysis of creole tense/aspect systems

As Dahl (1985:33) notes, attempts to identify cross-linguistic similarities in TMA systems can be based either on the sets of semantic features or dimensions out of which TMA systems can be built up, or on the actual TMA categories themselves. The former approach runs into several difficulties, not the least of which is deciding what the features or dimensions are, whether they are universal or not, and how many of them there are (*ibid.*). Dahl chooses a comparison based on categories rather than features, explaining as follows:

I shall suggest that the most salient ‘universals’ or, better, basic units of the general theory of TMA systems are rather atoms than elementary particles — i.e., categories rather than features. More concretely speaking, this means that I think of a language-specific category like, say, the English Perfect, as the

realization of a cross-linguistic category — or better, category-type — PERFECT, rather than as the realization of a set of features, say /+x, -y, +z/. (1985:33)

Henceforth I shall follow Dahl's practice of using capitals to refer to category types (e.g. PAST), initial capitals to refer to language-specific categories (e.g. the English Past), and lower-case letters with single quotation marks to refer to notional semantic categories (e.g. 'past'). Other scholars have adopted an approach quite similar to Dahl's. For instance, Anderson (1982) addresses the question of how to compare manifestations of the category-type PERFECT crosslinguistically. He suggests that it can be identified and compared on the basis of similarity on a language-independent map of semantic space. As Binnick (1991:100) explains, "it is precisely when certain concepts cluster and partially overlap, when categories in two languages cover largely the same portion of such a map, that we are justified in calling those categories by the same name". It is worth noting that both Anderson and Binnick acknowledge that there is usually 'partial overlap' between such categories. It would be rare and surprising to find that categories to which we give the same name across languages cover exactly the same area of semantic space, or display the same range of uses, even though they may be quite similar in their basic meaning and use.

There are certain other aspects of this approach to TMA systems that must be clarified before we proceed. In the first place, it is obvious that we need to distinguish between actual TMA categories which are grammaticized in a language, and the notional categories or areas of semantic space that may be expressed in various ways by different TMA categories, as well as by other means such as adverbials. For instance, the semantic space of 'past temporal reference' can be divided up in different ways and grammaticized differently from one language to another, or it can be conveyed by quite different TMA categories in the same language. An example of the former is the remoteness distinctions of past and future time reference which are grammaticized as distinct past and future tense categories in, say, many Bantu languages. An example of the latter is the fact that both Past tense and Perfect in English convey the sense of past temporal reference, though of course they differ in other respects.

Another consideration that is crucial to any attempt at cross-linguistic comparison is the fact that TMA categories in every language typically display a range of meanings and uses, i.e., interpretations in discourse. This is

why, as Binnick (1991:104) notes, there is a certain tension between the search for a unique meaning of a category (such as the Perfect) and the multiplicity of uses which all categories have. He suggests that “if the tenses (and aspects) have multiple uses, the effort to capture all the uses of a category under the umbrella of one categorial label is a difficult and perhaps futile one” (ibid.). A way out of this difficulty is offered by Dahl. He suggests that each TMA category has a ‘meaning’ that consists of several components, one or more of which is ‘basic’ or ‘dominant’ with regard to other ‘secondary’ meanings. Moreover, a TMA category may display different uses one or more of which is ‘primary’ or ‘prototypical’ by contrast with other ‘peripheral’ uses. Meaning and use (pragmatics) are thus closely related in this approach. Among other things, the notion of dominance provides a way of deciding when a particular category should be assigned status as a tense, modal or aspectual category, despite the fact that such a category may well combine ‘semantic parameters of temporal, aspectual or modal character’ (1985:23). For instance, Dahl notes that the use of the Future in English involves both ‘future time reference’ and modal notions such as ‘prediction’ and ‘intention.’ However, ‘the basic meaning, in the sense of dominant parameter, is ‘future time reference’ (1985:9-10). Hence it is possible to conclude that the Future in English is a tense rather than a modal category. Another case in point is the English Progressive. This category is often used to convey a sense of future intention, as in *I’m seeing her tonight*. But this is only a secondary or peripheral use, a pragmatic interpretation brought on by implicatures in the discourse context. The dominant meaning of Progressive, as seen in its primary uses, remains something like ‘ongoing activity.’

Failure to recognize such distinctions has sometimes led to disagreement or misunderstanding in discussions of creole TMA categories. A case in point is Spears’ (1990) criticism of Bickerton’s (1981) analysis of the HC auxiliary *ap* as a marker of ‘non-punctual’ aspect (subsuming progressive and habitual meanings). Spears bases his criticism on the fact that *ap* is not restricted to just these meanings, but is also used as a marker of futurity. (See also Boretzky 1983.) But this kind of secondary use of progressive markers is quite common, and does not affect the basic or primary meaning of the category. In general, then, the dominant meaning of a category is represented in its primary or prototypical uses, while secondary meanings tend to be associated with implicatures, that is, interpretations that can be inferred from contextual uses of the category. This brief explanation should suffice to make clear the assumptions

behind the present approach to creole TMA systems. As Dahl (1985:33) explains,

...the main criterion for identifying TMA categories cross-linguistically is by their foci or prototypical uses, and languages vary essentially in two respects: (i) which categories they choose out of the set of cross-linguistic categories and (ii) how they reduce the impreciseness that these categories have in choosing among the possible secondary or non-focal uses they have.

It should be clear by now that the approach I take here emphasizes the crucial role of the discourse context in deciding the interpretation of TMA categories. There are two aspects of the discourse context that are relevant. One is the actual linguistic context itself, for instance, the kinds of predicate as well as the types of temporal frame established within the sentential structure itself, e.g., by use of adverbials and the like. The other relates to the time-frame and points of temporal reference established in the wider discourse context. For example, in Western Caribbean creoles such as Jamaican creole and Belize creole, unmarked verbs may be interpreted as referring to present as well as past situations, and within these, to situations seen as occurring habitually as well as at specific times. These different interpretations depend crucially on the context established in the discourse. The same applies to various other TMA categories. For reasons like this, the methods of data collection employed in this study placed special emphasis on the need to capture as much discourse context as possible for the use of each tense/aspect category.

Another potential area of misunderstanding concerns the interaction between TMA categories and the inherent aspectual properties or *Aktionsarten* of different predicates. A case in point (mentioned earlier) is Bickerton's attempted explanation of the so-called 'non-punctual' aspectual category which he claims all radical creoles have. In the first place, there appears to be some confusion on B's part between the aspectual category itself and the *aktionsarten* of stative verbs. This leads him to treat non-punctuality as both a property of such verbs and as a manifestation of a TMA category. As a result, he erroneously claims that the 'non-punctual' category is compatible only with non-statives, since statives are by definition non-punctual. In fact, as several researchers have since pointed out, 'non-punctual' auxiliaries can occur with statives in creoles like GC and others, but then the meaning is 'habitual', whereas it may be interpreted as either 'progressive' or 'habitual' with non-statives. These facts may be explained more clearly by recognizing

the so-called 'non-punctual' category as in fact an Imperfective, with meanings that include both 'progressive' and 'habitual', with the former often 'neutralized' in stative contexts. As Dahl (1985:28) notes, the distinction between stative and non-stative constructions 'tends to be of considerable importance for TMA categories, in particular for aspectual categories, in that those categories tend to be less developed or wholly neutralized in stative contexts.' This is quite true of creole languages, including Sranan, as we shall see. Recognition of the need to keep the notion of a TMA category distinct from the 'aspectual potential' of verbal lexemes allows one to avoid 'a constant source of confusion in the study of TMA categories' (op. cit. 20).

3. Data and methodology

The data employed for this study were obtained in Suriname in the summer of 1994 and the winter of 1995. The data consist of two types: (a) elicitations from four native speakers based on a modified version of Dahl's (1985) TMA questionnaire, and (b) tape recordings of conversations among native speakers of African descent in working class areas of the capital city Paramaribo as well as rural areas in Para and Coronie.² All of the elicitations were done by me with informants who were bilingual in Sranan and English. The questionnaire used for the elicitations consisted of a number of sentences and short connected texts in English, which were offered for translation by informants. They were given clear indications, with additional explanation where necessary, of the contexts in which they were to envisage the sentences being uttered. For example, to elicit sentences containing a verb with past time reference, one of the prompts used is illustrated in (1):

- (1) [It is cold in the room. The window is closed.] Q: You OPEN the window (and closed it again)?

Material enclosed in square brackets is meant to provide a context for the utterance to be elicited (the translation of which is outside the brackets). Verbs are offered in bare form (capitalized in the text) so as to minimize the possibility of interference from English when translating. As Dahl (1985:45) notes, this approach meant that "the information necessary to choose the correct TMA category in the translation would have to be deduced from the sentence itself together with its context". Very often, in the course of elicitations

tions for this study, there were detailed discussions between the investigator and the informants to clarify the precise context and intended message for a given prompt sentence. Further discussion of the assumptions, methods and areas of difficulty associated with the use of the TMA questionnaire can be found in Dahl (1985:44-50).

The primary data for this study, however, came from recordings of actual conversations made by the principal fieldworker, Gladys Waterberg, a social worker with strong personal ties to the communities investigated, and herself a native speaker of Sranan. Most of the conversations involved at least two, sometimes several participants apart from the fieldworker. As a result, the speech recorded appears to be the natural everyday speech typical of interactions between close acquaintances. In other words, these data seem to represent the everyday vernacular Sranan of the informants.³ The recordings were transcribed, glossed and translated with the assistance of several native speakers, most of whom are employed for this purpose by the Summer Institute for Linguistics in Paramaribo.⁴

In the following discussion, example sentences elicited using the TMA questionnaire are identified by "E" plus the number of the prompt sentence. All other example sentences are from the recorded conversations and are identified by a tape number and page reference in parentheses.

4. Tense and aspect in Sranan

As already noted, the present analysis is restricted to tense and aspect categories, and will refer to modality only where it may be relevant to our understanding of the other two areas. (For a description of modality in SN, see Winford to appear.) Several researchers have attempted to describe the TMA system of SN or some part of it. Those who have attempted a comprehensive overview include Voorhoeve (1957, 1962), Seuren (1981, 1983) and Nickel & Wilner (1984). The last of these is quite sketchy, and repeats most of what the other studies say, so, I will say little more about it. Each of these analyses has contributed significantly to our understanding of TMA in SN, but none has given a complete, or completely accurate picture. Voorhoeve's (1957) analysis (repeated without significant change in Voorhoeve 1962) depicted a system consisting of three oppositions, in tense, mood and aspect, as illustrated in Table 2.

Table 2. Voorhoeve's (1957) analysis of TMA in Sranan

Mode	tempus			Aspect
	realis	past <i>ben</i>	present [zero]	
	non-realis	<i>ben sa</i>	<i>sa</i>	
	realis	<i>ben e</i>	<i>e</i>	
	non-realis	<i>ben sa e</i>	<i>sa e</i>	

This outline is remarkably similar to Bickerton's prototypical creole TMA system as illustrated earlier in Table 1, and it is arguable that B's prototype was in fact based on Voorhoeve's analysis of Sranan.⁵ This analysis is accurate in some respects, questionable in others, as will become apparent in the discussion of the individual categories below.

Seuren's (1981) analysis (repeated without significant change in Seuren 1983) offers a more comprehensive account of the TMA system of SN, including a detailed discussion of the system of modality, which I will not consider here. In addition, Seuren employs more conventional labels and terminology than those used by Voorhoeve. His analysis is summarized in Table 3:

Table 3. Seuren's (1981) TMA categories for Sranan

FORM	CATEGORY	MEANINGS
<i>Tense-aspect auxiliaries:</i>		
<i>ben</i>	Past Perfect	Past before past. 'Some speakers use it as a simple past.'
<i>o</i>	Future	'Future or temporal posteriority relative to a defined moment of time'
\emptyset	Past tense	Simple past; resultative perfect
<i>e</i>	Present tense	Durative; iterative ('A tense marker for the present or for temporal simultaneity')
<i>Modals:</i>		
<i>sa</i>	Uncertain Future	Potential or doubtful Future ('Future situations resulting from someone's [the speaker's] insistence, order, wish or promise'); other modal meanings
<i>kan/man</i>	Ability	'can', 'be able', 'may'
<i>musu</i>	Obligation	'must'

I take issue with some of Seuren's labelling and description of the other categories he postulates, as I will discuss when we come to the specific areas of the system in question.

5. A revised analysis of tense and aspect in Sranan

I argue here that the tense/aspect system of SN consists of the following categories;

Tense:	Relative Past.	<i>ben</i>
	Future	<i>o</i>
Aspect:	Imperfective	<i>e</i>
	Completive Perfect	<i>kba</i>
	Perfective	\emptyset (the unmarked verb)

These are the core categories to which the rest of this discussion will be devoted. In addition, however, the verb complex of SN contains several modal auxiliaries as well as various semi-auxiliaries and adverbial elements which convey various types of temporal, aspectual and modal meaning. A summary of the main forms and the meanings they express is provided in Table 4.

Table 4. A revised analysis of TMA in Sranan

FORM	CATEGORY	MEANINGS
<i>ben</i>	Relative past	Past events 'distanced' from S. Background past or 'framepast' especially in narratives. Past in relation to another reference point in the past.
<i>o</i>	Predictive Future	Later time reference; intention or prediction; predictability.
\emptyset	Perfective	States or events seen as unanalysed wholes.
<i>e</i>	Imperfective	Situations (both states and occurrences) seen as 'unbounded' and ongoing at reference time, which encompasses situations that are repeated, habitual, in progress or continuous
<i>kba</i>	Completive Perfect	Conveys the meaning 'already'. Expresses the sense of a 'perfect of result' with non-statives, and the sense of a state beginning in the past and continuing to the reference point with statives

Other aspectual markers

<i>kaba</i>	main verb	'finish'. Marker of terminative aspect with following verbs
<i>bigin</i>	main verb	'begin'. Marker of inceptive aspect with following verbs
<i>gwentí</i>	semi-aux (and noun)	Marker of customary behavior with following verb
<i>kon</i>	semi-aux.	Marker of 'achievements', 'realize a goal' 'come to V'
<i>du</i>	semi-aux	

Modal auxiliaries

<i>sa</i>	Uncertain Future	Uncertain or hypothetical future. Future situations (e.g. a wish or hope) not subject to speaker's control, or to the certainty of which the speaker is not committed
<i>musu</i>	Obligation	Deontic senses of 'must'. Marginal sense of necessity with certain (stative?) predicates
<i>kan</i>	Root Possibility	Deontic possibility, subject to moral or social law
<i>man</i>	Ability	(Deontic) ability subject to physical or natural law
<i>mag</i>	Permission	Deontic possibility imposed by authority (social, legal, etc.)

I will now examine the core tense/aspect categories individually.

6. The unmarked verb

It is convenient to begin this overview of tense/aspect categories in SN with the unmarked verb, which displays a wide range of uses and interpretations. Perhaps the best way to explain this is to say that the unmarked verb presents a situation as unanalyzed for any of the parameters of tense, aspect or mood. Hence it can lend itself to various interpretations in discourse, depending on the context and the predicate involved. It is debatable whether unmarked verbs should be characterized as 'marked with \emptyset ' as Bickerton (1975:46) seems to suggest. This would imply that \emptyset instantiates a specific TMA category with some basic semantic content, but that content would be extremely difficult to pin down, given the wide range of possible interpretations of unmarked verbs (see below). But it does seem plausible to claim that unmarked verbs are

perfective in aspect, as Comrie defines the notion. According to him, “Perfectivity indicates the view of a situation as a single whole, without regard to its internal structure” (1976:16). This is the so-called ‘totality’ view of ‘perfective’ and seems to be applicable in this case.⁶ In other words, the unmarked verb seems to be perfective in the sense that it presents a situation in its totality, without regard for its internal structure. In this respect, SN and other creoles are similar to languages like Abkhaz, Maung and others in which the perfective has zero expression and contrasts with a marked imperfective (Bybee et. al. 1994:90). I accept Bybee et al’s argument that ‘the semantic areas covered by zeroes are also systematic and comparable to those covered by overt grams’ and that ‘zeroes are attributed the same meanings as overt grams would have if they were available’ (1994:91). I therefore argue that zero-marked verbs in SN and other creoles instantiate a Perfective category.

The conventional wisdom is that the unmarked creole verb expresses present time reference with statives and past time reference with non-statives, but this is not completely accurate for SN or other Caribbean creoles. It is true, however, that once the discourse context establishes speech time (S) as the point of reference, the default interpretation of an unmarked stative is present, while that of an unmarked non-stative is past. The following examples from my data illustrate. The relevant verbs are in boldface:

- (2) a. *A pikin **wani** go sribi.*
 the child want go sleep
 ‘The child wants to go and sleep.’ (E: 2)
- b. *Mi **lobi** pley bal. Of a no bal, dan na volleybal.*
 I like play ball if COP NEG ball then COP volleyball
 *Den sortu sani dati mi **lobi**.*
 the-PL sort thing that I like
 ‘I like to play football (soccer). If not football, then it’s volleyball. Those sorts of things I like.’ (16A: 6)
- c. *A kamra kowru bikaa mi **opo** a fensre.*
 the room cold because I open the window
 ‘The room is cold because I opened the window.’ (E24)
- d. *Wan boi, ‘B’, a **dyuku** wan man **boro** en heri bere,*
 one boy B he stab one man cut.open his whole belly
 *ala en bere **kon** na doro.*
 all his belly come LOC out
 ‘A boy, ‘B’, he stabbed a guy and cut open his entire belly, all his guts came out.’ (16A: 18)

As is clear from the context, S is the point of reference in these examples. In such cases, the inherently dynamic character of non-statives, added to the perfective aspect conveyed by zero marking, favors their interpretation as past, as in examples c) and d). As Dahl notes, “a PFV verb will typically denote a single event, seen as an unanalyzed whole, with a well-defined result or end-state, located in the past” (1985:78). This description applies well to cases involving dynamic predicates. With statives, on the other hand, once the point of reference is S, the situation is interpreted as present, since such predicates describe unchanging situations, which the hearer will assume to be still in existence.

In other discourse contexts, however, unmarked verbs may take on quite different interpretations. For instance, once the discourse context establishes the reference point as past, unmarked statives can have past time reference. The following examples illustrate: (Statives with past reference are underlined and Dutch words which are clear code-switches are italicized.)⁷

- (3) a. *Mi tyari den sani gi a man; a man teki a*
 I carry the-PL thing give the man the man take the
zaak op. Ma tegelijkertijd a man abi a zaak
 case up but at.the.same.time the man have the case
fu a tra man di kiri a boi.
 of the other man who kill the boy
 ‘I carried all those documents for the man; he took the case. But
 at the same time, the guy had the case of the other man who
 killed my son.’ (39A: 25)
- b. *Dus mi ben go a wan tu suma kaba.*
 so I PAST go LOC one two person COMP
Dan mi prakseri kon mi kon na oom N.
 then I think come I come LOC uncle N.
 ‘So I already went to one or two people. Then I thought, let me
 come to uncle N.’ (39A: 30)
- c. *A bigi suma kon dape dan a si mi. Dan a*
 the big person come there then she see me then she
lobi mi. Omdat a so pikin mi de a e si mi
 like me because COP so small I be she IMP see me
e wroko.
 IMP work
 ‘The elderly lady came there and saw me. And she liked me.
 Because I was so little and she saw how I worked.’ (34B: 21)

Similar uses of unmarked statives are to be found in other creoles such as GC (Jaganauth 1988; Winford 1993), JC (Winford 1993) and Belize Creole (Winford 1994). These facts call into question the conventional wisdom on the use of unmarked verbs.

In addition, there are uses of unmarked verbs which convey a kind of perfect, as mentioned by Voorhoeve and Seuren. The following are examples.

- (4) a. A: *Ma unu no si a man dya a Coronie?*
 But you NEG see the man here LOC Coronie
 B: *Persoonlijk noyti mi no si en dya a Coronie.*
 Personally never I NEG see him here LOC Coronie
 ‘But you haven’t seen the guy here in Coronie?’
 Personally, I’ve never seen him here in Coronie.’ (39A:04)
- b. A: *Ayi, unu go na en?*
 Yes you go LOC him
 B: *Neen, mi no go ete. Te mi e drai...*
 No I NEG go yet when I IMP return
 ‘Yes, did you go to him?’
 ‘No, I haven’t gone yet. When I’m going back....’ (39A:10)

The sense of a ‘perfect’ conveyed in these examples is conveyed partly by the discourse context and partly by adverbials like *noyti* ‘never’ and *ete* ‘yet’. But this is a secondary meaning of the unmarked verb. In general, unmarked verbs tend to refer to situations that have current relevance — a notion closely associated with typical uses of PERFECT. But it would be inaccurate to claim that the basic or dominant meaning conveyed by unmarked verbs is that of a perfect. The use of such verbs may be better understood when we compare them to the use of Relative Past *ben*, to be discussed below.

Finally, the neutral temporal character of unmarked verbs allows them to be used freely in the protasis (*if*-clause) of conditionals to express future possibility.

- (5) a. *En efu a man dati wini en, a kondre e kon*
 and if the man that win it the country IMP come
 bun zeker.
 good surely
 ‘And if that man wins it [the election], the country will surely improve.’ (16A: 29)

- b. *Efu pikinso wan druppel nomo kisi yu skin,*
 if little-so one drop only catch your skin
yu skin e bron.
 your skin IMP burn
 ‘If even so much as a drop touches your skin, it burns.’ (16A: 22)

Again, these uses have parallels in other creoles.⁸

7. Tense in Sranan

Comrie’s (1985) well-known characterization of tense as “grammaticalized location in time” is a good starting point for our discussion of tense categories in SN. As Chung & Timberlake (1985: 203) explain, “an event occurs on an interval of time, the event frame. Tense locates the event in time by comparing the position of the frame with respect to the tense locus”. Cross-linguistically, there is a broad distinction between tense systems in which the tense locus or point of reference is S (absolute tense systems), and those in which it may (also) be some other point in time (relative tense systems). I will argue here that SN, like many other creoles, has a relative tense system in which the tense locus may be either S or some other reference point. I shall also show that there are only two tense categories in SN, (Relative) Past and Future, though of course past and future temporal meaning can be conveyed by other means. There is no present tense category in SN. As in other creoles, present time reference is conveyed by categories that are primarily aspectual in character, such as Imperfective, as in SN, or, in other creoles, Progressive and Present Habitual. Most of the following discussion will be concerned with the expression of tense in main clauses. The way time reference is expressed in subordinate clauses will be touched on briefly later.

7.1. *Relative past tense in Sranan*

The Relative Past tense category of SN is instantiated by the auxiliary *ben*. Following Bickerton (1981), many creolists refer to this and similar past tense categories in other creoles by the label ‘anterior’ — a term which has other uses in the literature.⁹ Bickerton’s view that this category expresses ‘very roughly past before past for action verbs and past for stative verbs’ has also

gained wide acceptance, and has led no less a scholar than Dahl (1985: 119) to (mis-)interpret the creole category as a unique instance of what he calls “past restricted to stative contexts,” and to suggest that it is “important to the general theory of TMA categories — and maybe linguistic theory in general — since it appears to be fairly common in creole languages as noted in Bickerton (1981) and elsewhere.” In fact there is nothing unique about this category, which parallels relative past categories in other languages. It turns out that Bickerton’s (1981) characterization of ‘anterior’ is quite inaccurate for SN as well as other creoles which have been adequately studied. A much better analysis is in fact offered in Bickerton’s earlier (1975, 1979) accounts of the Past tense auxiliary *bin* in GC and other creoles. He notes that “a [+anterior] action does not have to be a ‘past before past’, since it could be regarded as both related and prior to a state of affairs at present in existence” (1975:46). In other words, some uses of the past category are equivalent to an absolute past, while others are equivalent to a ‘past before past.’ Bickerton’s earlier insight was developed further by Jaganauth (1987, 1988), who analyzed GC *bin* as a Relative Past marker. Winford (1993) provided further evidence in support of this analysis. I argue here that the function of SN *ben*, like GC *bin*, is to locate a situation as past in relation to some other point or interval of time (the tense locus) which may be either the moment of speech or some reference point in the past.

This description corresponds quite closely to Comrie’s characterization of a relative tense as one which locates a situation “at, before or after a reference point given by the context” (1985:65). As Comrie (p. 63) makes clear, such a reference point may include the moment of speech. As we shall see, however, it is not always easy to identify the point in time given by the context (p. 56). In particular, there are problems with his assumption that relative tenses always derive their interpretation from other verbs in the discourse, though he adds the important qualification “barring contextual indications to the contrary” (p. 58). As we shall see, such indications can be crucial.

My interpretation of this tense category in both GC and Sranan is in keeping with Bickerton’s observation that the meaning of GC *bin* shifts between ‘simple past’, ‘remote past’ and ‘past before ‘past.’ As Jaganauth (1988) puts it, ‘if we bear in mind that it is the reference point which moves,’ then there is no need to view the various uses of *bin* as distinct. This analysis has been shown to apply equally well to JC (*ben*) and its counterparts in

Caribbean English creoles (henceforth CEC) generally (Winford 1993). Similar conclusions have been reached in the discourse-based accounts offered by Pollard (1989) for Jamaican Creole, Spears (1993) for Haitian Creole and Wilner (1992) for Sranan.

We can conclude therefore that the meaning of SN *ben* is simply relative past. The semantics of this auxiliary allow it to be used in various interpretations in actual discourse. These pragmatic uses are best illustrated by examples which include discourse context, rather than by single sentences extracted from their context. Reliance on the latter kind of data has been responsible for much of the misunderstanding about this and other creole TMA categories.

The prototypical use of *ben* is to locate some situation as occurring prior to the reference point under focus in the discourse. The effect of this is to distance the situation from the reference point, and “background” it in some way. The following elicited examples (with the context in square brackets) illustrate:

- (6) a. [Q: Did you ever meet my brother, who is now dead?]
Iya, mi ben miti nanga en wan leisi.
 yes I PAST meet with him one time
 (?? *Iya, mi miti nanga en wan leisi.*)
 ‘Yes, I met him once.’ (E-32)
- b. [Q: Did you open the window (and close it again)?
 (The window is now closed).]
Aay, na mi ben opo a fensre.
 yes COP I PAST open the window
 (?? *Aay, na mi opo a fensre.*)
 ‘Yes, it was me that opened the window.’ (E-35)

In each of these cases, the reference point established in the discourse is itself a past event (the death of the brother, and the closing of the window). In these cases, *ben* situates the event it marks as prior to these reference points.

Better illustration of the use of *ben*, however, comes from natural discourse, most instructively when we contrast uses of *ben* with unmarked verbs in extended discourse. Narratives in particular provide good illustration of the prototypical uses of *ben*. The following narrative, separated into parts (a-f) for convenience, illustrates (Instances of zero marking and *ben* are in boldface for easier comparison):

- (7) a. *Ma wan vyftien yari, wan twintig yari pasa tu,*
 but one fifteen year one twenty year pass too
*wan man K. **ben** dede a Coppename.*
 one man K PAST die LOC Coppename
- b. *Leti so wan sortu fasi, tog. A \emptyset gwe ini a busi,*
 right so one sort way TAG he go in the bush
an no \emptyset kon noyti moro a dorosei.
 and NEG come never more LOC outside
- c. *Ma a man dati, sani **ben** musu fu miti a man*
 but the man that thing PAST must to meet the man
*dati. Omdat a man dati **ben** abi tapu na en skin.*
 that because the man that PAST have obeah LOC his skin
- d. *A man, dus, a man **ben** kiri kapoewa,*
 the man so the man PAST kill kapoewa
*ala meti san kon a **ben** kiri.*
 all meat thing come he PAST kill
- e. *Dus a ten \emptyset kon now taki a san \emptyset drai onderste*
 so the time come now that the thing turn upside
boven gi en. A \emptyset go wan leisi a busi, noyti moro a \emptyset
 down for him he go one time LOC bush never more he
kon a syoro.
 come LOC shore
- f. *Dati kan de so wan twintig, kande wan tertig yari tori.*
 that can be so one twenty perhaps one thirty year story
 (Tape 39A: 13-14).

‘But some fifteen to twenty years ago, a man by the name of K. died at Coppename. In a similar fashion, right? He went into the woods and never came out again. But things were bound to happen to that man. Because he had obeah on his body. The man, so, the man had killed kapoewa [a wild pig], all kinds of animals. So now the time came when the tables were turned on him. He went into the woods one time, and never returned. That might have been some twenty, perhaps thirty years ago.’

In the above extract, the use of *ben* in the opening sentence (7a) places the situation in the remote past, distancing it from the moment of speech. The next sentence (7b) introduces the story line, employing unmarked verbs to convey the main events of the narrative. The following sentences (7c-d) provide

background information to explain the cause of these events, and once again *ben* is used to establish this background. The final sentences (7e) return to the main story line, which calls once more for the use of unmarked verbs. Sentence (7f) provides the closing comment about the time of the incident, and the unmarked verb has present time reference. The following extract (Tape 39A: 23-24) reveals a similar pattern of use.

- (8) A. *Ma fa a pikin fu Oom N Ø du dede dan?*
 but how the child of uncle N do die then
 'But how did Uncle N's child actually die, then?'

- B. *Ayi, a wan moi vraag, noh. Wel, mi ben de a*
 yes COP one nice question TAG well I PAST be LOC
Novar. Mi Ø de a Novar, mi Ø de nanga wan man,
 Novar I be LOC Novar I be with one man
ben (e) taki tori nanga wan man.
 PAST IMP talk story with one man
 'Yes, it's a good question, no? Well, I was at Novar. I was at
 Novar, I was with a guy, I was talking to a guy.'

Dan wan man Ø ry kon nanga wan bromfiets, ma a
 then one man ride come with one bike but the
man di Ø kon teki mi, en nanga den suma dati no Ø
 man REL come take me he and the-PL people DEM NEG
de bun. Dan a man Ø tan a strati, dan a e
 COP good then the man stay LOC street then he IMP
bari kari mi 'D, D, Oom N, Oom N, Oom N'.
 shout call me D, D uncle N. uncle N. uncle N.

'Then a guy came on a moped, but the guy who came to pick me up, he and those people (I was talking to) were not on speaking terms. So the guy stayed in the street' then he called me 'D, D, Uncle N, Uncle N, Uncle N'

Dan mi e taki taki a man Ø abi mi fanowdu,
 then I IMP talk COMP the man have me need
san a man no e kon.
 why the man NEG IMP come
Mi no e go. Okay, dan a man, a man di mi
 I NEG IMP go. okay, then the man the man REL I

nanga en e taki, a man Ø taki, yu kan go want
 and he IMP talk, the man say you can go because
a man no e kon. A no o kon dya.
 the man NEG IMP come. he NEG FUT come here
 ‘So I was saying that the guy needed me, why wasn’t the guy
 coming to me? I’m not going. Okay, then the guy, the guy I was
 talking to, the guy said, you can go for he’s not going to come.
 He won’t come here.’

Di mi Ø go a man Ø taigi mi taki taki ‘Oom N,
 when I go the man tell me say say uncle N
sidon a baka a fiets. Oto Ø naki a boi fu yu.’ Mi Ø
 sit LOC back the bike car hit the boy of you I
taki ‘san! oto Ø naki en.’ Mi Ø taki efu oto naki en a
 say what car hit him I say if car hit him he
o dede. Merkwaaardig....
 FUT die Oddly

‘When I went, the guy told me ‘Uncle N, sit on the back of the
 bike. A car has knocked down your son.’ I said ‘what! a car
 knocked him down?’ I said if a car has knocked him down he’ll
 die. Oddly enough....

- A. *Oh, Oom N Ø taki a sani dati?*
 oh uncle N. say the thing DEM
 ‘Oh, Uncle N said that?’
- B. *Iya, mi Ø taki efu oto Ø naki en a o dede.*
 yes I say if car hit him he FUT die
Merkwaaardig di unu Ø go, a man Ø hari a laaste
 Oddly when we go the man blow his last
blo ‘haah’, klari.
 breath haa clear
 ‘Yes, I said if he was struck by a car he’ll die. Oddly enough,
 when we went, he took his final breath ‘haah’, that was it.’
- A. *San, leti fesi Oom N?*
 what right face uncle N
 ‘What! right before Uncle N?’
- B. *Iya, a laaste blo ‘haah’, klari. Mi Ø taki a o dede.*
 yes the last breath haah clear I say he FUT die
 ‘Yes, his final breath ‘haah’, that was it. I said he would die.’

- A. *Dan omeni yari a ben abi?*
 then how.many year he PAST have
 ‘So how old was he?’
- B. *Uh, omeni yari P. ben sa abi? Vyfen twintig, noh?*
 uh how.many year P. PASTFUT have twenty five no
 ‘Uh, how old would P. have been, twenty-five, no?’
- C. *A ben pasa wan vyf en twintig so, iya.*
 he PAST pass one twenty and five so yes
 ‘He was over twenty-five or so, yes.’

Notice how *ben* is used in the opening sentences of B’s story to establish the backdrop for his narrative concerning his son’s death. The first *de* is marked by past, but the next two are not, since the time reference of this verb is already established. However, *ben* is repeated with the new verb *taki*, to reaffirm that this is background information. The story line then begins, and continues almost to the end of this abstract. All verbs are unmarked in this part of the narrative. In the final three sentences, *ben* is once more used to convey the sense of a past situation that is distanced from current relevance, that is, the son’s age before or at the time of his death.

In general, then, *ben* marks situations that provide a kind of backdrop against which the main storyline develops. The events that make up the main narrative are conveyed by unmarked verbs, whose time reference is already established as past by the context, including *ben* itself. As can be seen, both statives and non-statives may be either unmarked (if they are part of the storyline) or marked by *ben* (if they are part of the background for the story).¹⁰ Pollard’s (1989:63) explanation of the uses of *ben* vs unmarked verbs in JC narratives applies equally well to Sranan.

Information that is more highly relevant as a response to a given question, that is salient or in-focus when compared with other information, is Foreground information and is expressed by the unmarked form of the verb. Information which is perhaps necessary but less highly relevant, less salient, less in immediate focus is Background information and is expressed by the verb plus particle *en*.

Wilner (1992:4) agrees that “the distinction between foreground and background is essential” in understanding the use of *ben* in SN, and provides more detailed analysis of this use in several SN narratives.

However, the distinction between foreground and background is not always as simple as one might assume.¹¹ The uses of *ben* discussed so far

demonstrate clearly how the past marker locates situations prior to some reference point established in the discourse. Identifying the reference point, however, is not always so straightforward. It is not the case, for instance, that the reference point is always established by the time reference of another verb in the same or an adjacent sentence. There are many examples in my data involving sequences of verbs in the same or consecutive sentences where *ben* is not used, even though it seems clear that one event preceded another. The following are examples:

- (9) a. *Di mi doro na oso esde, mi bradi gi*
 when I arrive LOC house yesterday my brother give
mi moni
 me money
 ‘When I reached home yesterday, my brother gave me money.’
 (E: HL 29)
- b. *Baka di a dringi a dresi, a koso wan*
 after when he drink the medicine he cough one
heri yuru.
 whole hour
 ‘After he drank the medicine, he coughed for a whole hour.’ (E:
 HL 27)

The likely explanation in these cases is that the speaker regards both events as absolute pasts, i.e., located as past relative to S, and sees no need to distance one from the other. There are also sentences which describe two events in sequence, in which *ben* is used in the same way as the pluperfect of English, that is, to locate one of the past events as prior to the other.

- (10) *Di mi doro na oso esde, mi brada **ben** gwe*
 when I arrive LOC house yesterday my brother PAST leave
kaba.
 COMP
 ‘When I reached home yesterday, my brother had already left.’ (E:
 HL 71)

But *ben* does not always take its reference point from a verb in the same sentence. One example is the following:

- (11) A. a. *Ma dan pe a boi dati de nownow? A no de?*
 but then where the boy DEM COP now he NEG COP
 'But then where's that boy now? He's not around?'
- B. b. *Suma, Bra? Mi no si en esde.*
 who Bra I NEG see him yesterday
 'Who? Bra? I didn't see him yesterday.'
- C. c. *Mi si en ini a wiki disi*
 I see him in the week this
 'I saw him this week'
- B. d. *Traesde mi miti en even.*
 other.yesterday I meet him briefly
A ben taigi mi a o kon na fesisey baka.
 he PAST tell me he FUT come LOC front.side back
 'Day before yesterday I met him briefly. He told me he
 would come to the front again.'
- e. *A bo¹² go teki Remes na Albina mi bribi.*
 he PAST+FUT go take Remes LOC Albina I believe
 'He was going to pick up Remes in Albina, I think.'
- f. *Mi no sabi efu a go ete.*
 I NEG know if he go yet
 'I don't know if he's gone yet.' (8A: 24)

In this case, the unmarked verbs in b) and c) take S as their reference point, since these sentences are replies to the question in a). This is also true of the unmarked verb in the first sentence in d). Why, then, is the verb in the following sentence marked with *ben*, since it is clear that the event in the previous sentence precedes the event marked by *ben*? The reason is simply that the speaker wishes to background this information in relation to his present knowledge of Bra's whereabouts, as stated in the following sentence f).

In short, the reference point for an event marked by *ben* is not necessarily another past event in the same sentence or preceding discourse. Use of *ben* has more to do with the speaker's wish to distance that event from some other situation which s/he wishes to foreground. This is why we often find *ben* and unmarked verbs alternating throughout the course of a narrative, so that it sometimes seems impossible to predict when one or the other will be used. The following extract of a conversation between three teenagers, 'G', 'F' and 'J' illustrates (Tape 16A: 30-31):

- (12) J. *Ma a no wan fasi mi no kan feni wan valies*
 but COP NEG one way I NEG can find one suitcase
moni nownow de nanga soso Amerikan nanga
 money now there with only American and
Holland moni yungu.
 Dutch money man
 'But if there was a way I could find a suitcase of money right
 now with only American and Dutch money, man!'
- F. *Yu, wan lo sprookjes onder water dati yu e wani*
 you one lot fairy tales under water that you IMP want
ferteri.
 tell
 'You, it's a whole bunch of fairy tales you're always wanting to
 tell'
- J. *A no sprookjes onder water, yere. Na a srefi stori.*
 COP NEG fairy tales under water hear COP the self story
Leki wan mati fu mi san mi ben sabi. A man Ø
 like one friend of me REL I PAST know the man
gwe na Holland. H.
 go LOC Holland H.
 'It's not fairy tales, okay? It's a true story. Like a friend of mine
 whom I knew. The guy's gone to Holland, H.'
- F. *Ayi. 'Yeah'*
- J. *A man dati ben feni volgens mij abra*
 the man DEM PAST find according.to me over
dusun Hollands.
 thousand Dutch
 'That guy found, I believe, over one thousand Dutch (guil-
 ders).'
- G. *Wakti, suma yu e taki? Babu fu dya bakasei?*
 wait who you IMP talk Hindustani from here backside
 'Hold on, who are you talking about? The Hindustani from
 behind here?'
- J. *Ayi, a man fu dya bakasey a mi.*
 yes the man from here backside LOC me
A man ben feni wan dusun Hollands.
 the man PAST find one thousand Dutch (guilders)

Dan san a man Ø baka, na echt a man Ø baka,
 then what the man bake COP real the man bake
yere. A man Ø bai wan video, wan Nintendo game,
 hear. the man buy one video one nintendo game
srefi. A man ben bai wan tv gi en mama, yu
 self the man PAST buy one tv give his mama you
sabi.

know

‘Yes, the guy here behind where I live. The guy found a thousand Dutch guilders. And the guy did well for himself [*baka* is slang]. The guy bought a VCR, a nintendo game even. The guy bought a tv for his mother, you know.’

F. *Huh?*

J. *Yu sabi omeni sani a man Ø du. En a man Ø*
 you know how many thing the man do and the man
taki te a man go a Holland dan a man o
 say when the man go LOC Holland then the man FUT
seni wan sani gi mi.

send one thing give me

‘You know how many things the guy did? And he said when he went to Holland he would send something for me.’ (Tape 16A-30-31).

The pattern of use of *ben* vs unmarked verbs in the above extract may appear quite random at first. But it becomes clearer if we understand that the speaker’s narrative contains two story-lines, one of which — that related to H.’s finding and past spending of the money — provides a backdrop to the other — H.’s promise to buy something for the speaker himself. Events related to the first storyline are distanced by *ben*, while those related to the promise to send *wan sani gi mi* are brought in focus by unmarked verbs, which convey a sense of current relevance. Without this understanding, the choice of *ben* or an unmarked verb would be quite difficult to predict, since it depends so much on the perspective from which the speaker wishes to view and present the situation (Cf. Reichenbach’s [1947] ‘temporal point of view.’)¹³ In short, it would be simplistic to expect that the reference point for the situation marked by *ben* is always established by the time reference of other verbs in the immediate discourse. Instead, it would appear that it is the speaker’s own

perception of temporal relationships in some wider context that triggers the selection of *ben*.

It is clear from the above examples that *ben* and the unmarked verb convey quite different senses of past reference. This calls into question the claims made by Seuren (1981) and Sankoff (1990) that *ben* and the unmarked verb are in competition with each other. Seuren (1981:1050) goes so far as to suggest that *ben* is an older form that is giving way to the unmarked verb. As he puts it, "In older forms of the language, one often finds *mi ben nyam* ['I ate'] next to the frequent *mi nyam*. The form *mi ben nyam* is still not extinct or obsolete, though apparently less frequent than it used to be". The fact is that *ben* is far from becoming obsolete or less frequent in SN. It is a fully productive part of the tense/aspect system. This also contradicts Sankoff's claim that in SN "the category [anterior] has not fully grammaticized formally, such that unmarked forms can be considered to have a particular instantiation as zero-marked" (1990:309-10).

One final function of the Relative Past marker which must be mentioned is its use in conditional clauses. In SN, as in other creoles, *ben* is used in the protasis of conditions to convey counterfactuality, and is also used in combination with future markers *o* or *sa* in the consequent (apodosis) to convey past or present counterfactuality or probable future. The following examples illustrate.

- (13) a. *Efu mi ben de datra, mi ben sa/ bo*
 if I PAST COP doctor I PAST FUT PAST+FUT
abi furu moni.
 have full money
 'If I were a doctor, I might/would have a lot of money' (E-HL)
- b. *Efu mi ben gudu, mi ben sa/ bo bai wan oto.*
 if I PAST rich I PAST FUT PAST+FUT buy one car
 'If I were rich, I might/would buy a car.' (E-HL)
- c. *Efu mi ben de unu, yere, mi no bo luku*
 if I PAST COP you hear I NEG PAST+FUT look
vriendin ete.
 girlfriend yet
 'If I were you, right?, I wouldn't look for a girlfriend yet.'
 (16B:12)
- d. *Efu unu ben tenapu luku, unu ben sa bori en*
 if you PAST standup look you PAST FUT cook it
leki mi.
 like me

‘If you had stood and looked, you’d probably cook it [a soup] like me.’ (1A: 08)

Fuller discussion of conditionals in SN can be found in Winford (to appear). Related to the above are other uses which Wilner (1994) has referred to as ‘non-temporal.’ He discusses various examples in which *ben* is used to convey a sense of hypotheticality or suggestion, to ‘soften’ the effect of a request, an invitation, or a piece of advice. The following examples illustrate (example c) is from Wilner):

- (14) a. *Oh, ma dan a ben kan kom taki now, tog?*
 oh, but then she PAST can come talk now TAG
 ‘Oh, but then she could come and talk now, right?’ (39A: 38)
- b. *Mi ben sa wani meki yu ben sa taki en tapu*
 I PAST FUT want make you PAST FUT talk it top
a tape.
 the tape
 ‘I would like to ask if you could tell it on tape.’ (15A: line 125)
- c. *Ma mi ben sa gi yu a rai fu yu no teki*
 but I PAST FUT give you an advice for you NEG take
en baka, yere?
 him back hear
 ‘But I would advise you not to take her back, hear?’

In these cases, *ben* adds an element of indirectness to the request or invitation. As Wilner points out, “There is no assertion of reality; rather, there are attitudes or feelings or a judgment being communicated. This is traditionally not the realm of ‘tense’ but of ‘mood’” (1994:2).

7.2. Future tense in Sranan

As we noted earlier, Bickerton’s (1981) prototypical TMA system included a category of ‘irrealis’ mood which was claimed to subsume future time reference as well as various modal notions. It has in fact become the tendency among creolists to refer to any category which expresses futurity in a creole as ‘irrealis.’ The label is in fact inappropriate, since no creole has a single TMA category which expresses all possible irrealis meanings. As Bickerton himself notes, irrealis meaning includes “all states and actions which have not actually occurred, whether these are expressed by future or conditional tenses or by

modals" (1975:42). The fact is that all creoles distinguish future tense categories from others which express different types of modal meaning, and also have distinctive strategies for conveying conditional meaning.¹⁴

I will demonstrate here that SN has two categories whose primary function is to express futurity. One of these conveys the sense of a relatively certain or predictive future, and is expressed by auxiliary *o* (< *go*). The other conveys the sense of a dubitative or uncertain future, and is expressed by auxiliary *sa*. I will argue here, following Seuren (1981), Nickel & Wilner (1984) and Wendelaar & Koefoed (1988) that *o* is a marker of future tense, while *sa* is more modal in character, expressing uncertain future and related meanings. Voorhoeve (1957) offers a rather different analysis. He notes that 'it is possible to render the occidental future in Sranan by means of *sa* + verb and by means of *e-go* + verb', but argues that *e-go* is 'a verbal group consisting of the auxiliary *e* plus main verb *go*, and hence cannot be described as a category within the system of verbal forms (formed by prefix + verb)' (1957:379). Voorhoeve's analysis of *e-go* is essentially correct. I shall show that it can express the sense of a prospective future, and does indeed consist of Imperfective *e* plus main-verbal *go*, which takes a VP complement. However, Voorhoeve fails to distinguish *e-go* from *o*, which is a true future tense marker, and he therefore erroneously excludes the latter from his TMA system for SN. He also errs in rejecting earlier arguments by Simons (1954), Donicie (1954) and others that SN has two types of future. For these reasons, he concluded that *sa* and not *o* belonged in the TMA system of SN. Voorhoeve's claim that *sa* conveyed the meaning 'non-realis', combining both futurity and certain modal notions, was apparently carried over into Bickerton's claim that 'radical' creoles have an 'irrealis' category with a similar range of meanings.

The failure to distinguish clearly between future tense categories and modal ones reflects a common problem faced by analysts of TMA systems. As Dahl (1985:103) points out, 'the distinction between tense and mood becomes blurred when it comes to the future.' One reason he suggests is that

Normally, when we talk about the future, we are either talking about someone's plans, intentions or obligations, or we are making a prediction or extrapolation from the present state of the world. As a direct consequence, a sentence that refers to the future will almost always differ modally from a sentence with non-future time reference. (ibid.)

Chung & Timberlake (1985:243) on the other hand, argue that future morphemes tend to have modal overtones because modality involves degrees

of uncertainty, and “situations in the future are inherently uncertain as to actuality”. All of these explanations are no doubt relevant. At any rate, the problem of how to distinguish a future tense from a mood category remains. A way out of this difficulty is offered by Dahl’s distinction between the dominant and secondary meanings (uses) of a category. In his approach, as we have seen, the dominant meaning is represented in the prototypical or ‘focal’ uses of the category, while secondary meanings tend to be associated with implicatures inferred from the context, and are not regarded as belonging to its proper meaning. For example, the modal features of Future tense in English and other languages have led some linguists to claim it is a mood rather than a tense. Yet, as Dahl argues, ‘future time reference’ is a more constant element of FUTURE than the modal features of this category-type. Hence future time reference must be regarded as the dominant feature of FUTURE and the traditional view of Future as a tense can thus be defended’ (1985:106-7). I argue here that this is precisely true of Future *o* in SN. By contrast, the auxiliary *sa* must be analyzed as a modal auxiliary which expresses a dubitative or hypothetical future, along with associated epistemic notions. Voorhoeve’s description of *sa* is essentially correct. He notes that “Sranan uses the *sa*- form rather in order to express a wish, an intention or an expectation of the subject than a future tense” (1957:380). In short, *sa* is used for events conceived of as potentially possible but uncertain.

Seuren (1981) has suggested that *sa* and *o* are in competition and that the latter is replacing the former. I find no evidence of this in my data, which show that *sa* and *o* express quite different meanings, i.e., quite different degrees of commitment to the certainty of the event in question. The difference is also quite clear to my informants. As one of them (HL) puts it, “*o* is more certain; *sa* means you’re not sure”. Wendelaar & Koefoed (1988) also find a clear distinction between the two, based on elicited data and judgments of appropriate usage. I am in basic agreement with their view that, in general, “*o* is used for (a) the relatively certain realizable future and (b) the (absolute) certain non-realizable future; *sa* is used for (a) the uncertain (doubtful) realizable future and (b) the neutral and uncertain non-realizable future” (1988:69) [my translation].¹⁵

The following examples illustrate the use of *o* in its primary function as a predictive future.

- (15) a. *Efu yu no wroko, dan you no o nyan, tog.*
 if you NEG work then you NEG FUT eat TAG
 ‘If you don’t work, then you won’t eat, right?’ (34B-5)

- b. *Pas te unu kaba nanga skoro dan wi o meki*
 only when we finish with school then we FUT make
pikin nanga den sani dati.
 child and the-PL thing DEM.
 ‘Only when we finish with school, then we’ll have kids and all those things.’ (16B: 12)
- c. *Unu train tranga en mi o sori den suma na*
 we train hard and I FUT show the-PL person COP
mi, ayi. Mi echt. Mi o dunk den bal ernstig
 I yes. I real I FUT dunk the-PL ball seriously
tide, ay.
 today yes
 ‘We practised hard and I’ll show those guys who I am, yes. I’m for real. I’ll dunk the ball seriously today, yes sir.’ (16A: 6)

Note the modal overtones of intention in cases where the speaker is the agent, and refers to events that are under his/her control, as in example (c) above. On the other hand, *sa* is used where the speaker has little commitment to the likelihood of the future event, as in the following examples.

- (16) a. A: *San yu denki musu pasa taki Coronie kon*
 what you think must pass COMP Coronie come
kisi libi baka?
 catch life back
 B: *A sa kisi en? Mi no abi a bribi dati.*
 it FUT catch it I NEG have the belief DEM
 ‘What do you think should happen for Coronie to become alive again?’
 ‘Will it? I don’t believe so.’ (49B: 52)
- b. *Dan te mi miti en mi sa aksi en. Mi sa*
 then when I meet him I FUT ask him I FUT
taigi en mi kon na en mama na Coronie.
 tell him I come LOC his mother LOC Coronie
 ‘Then when I see him I will ask him. I’ll tell him I visited his mother at Coronie.’ (34A-13)
- c. *Neen, mi taki en, mi taki efu a kon dan mi*
 no I say him I say if it come then my

granpikin nanga mi afopikin sa kon
 grandchildren and my great-grandchildren FUT come
meti en. Ma mi no o si en moro.
 meet it but I NEG FUT see it more
 ‘No, I told him I said that if it comes then my grandchildren
 and great grandchildren may see it, but I won’t see it again.’
 (49B:49)

The fact that both *o* and *sa* refer to the future means that they often appear to be used interchangeably, but in fact they convey quite different senses, and cannot be said to be truly in competition. This can be seen in examples like the following:

- (17) a. *Efu mi kisi furu moni, mi o/sa bai wan bigi oso.*
 if I get full money I FUT/FUT buy one big house
 ‘If I get a lot of money, I will/may buy a big house.’ (E-56)
- b. *A yuru te yu kon baka mi o/sa kba skriŋi*
 the hour when you come back I FUT/FUT finish write
a briŋi disi.
 the letter DEM
 ‘By the time you come back, I shall/should have finished
 writing this letter.’ (E-66)
- c. *Uhum. ‘Meri Mi’ betekent dati a grontapu o kon*
 uhum meri mi mean that the world FUT come
bun baka en wi sa libi bun baka nanga
 good back and we FUT live good back with
alasuma. En mi lobi a pisi dati.
 everyone and I love the piece DEM
 ‘Uhum. ‘Meri Mi’ (‘Mess with me’) means that the world will
 become good again and we will all live well again with every-
 one. And I like that song.’ (Tape 16A-28)

Another primary use of *sa* is to express hypothetical meanings, the sense of English *would* or *could*. It is especially frequent in formulaic expressions like *San yu sa du* ‘What can you do?’ or *Fa mi sa taki* ‘What should I say’. The following examples illustrate:

- (18) a. *Mi ben de pikin nengre, mi sa taki den tien, elf.*
 I PAST COP small child I FUT say the-PL ten eleven

Dan mi lowe.

then I run.away

‘I was a small child, I would say ten or eleven, and I ran away.’

- b. *Ma dan tog, san yu sa taki? Pe wan pikin*
but then TAG what you FUT say where one small
nengre de, zegen nanga blesi de ini a
child COP blessings and blessings COP in the
oso dati.

house DEM

‘But then, what can you say? Where there’s a child, there are also blessings in that house.’ (49A:19)

Future *o* cannot be used in cases like these, since it lacks the strong modal overtones of uncertainty associated with *sa*. Quite similar to these are cases in which *sa* is used to convey a wish or hope of the speaker:

- (19) a. *Ma, uh, mi sa aksi mevrouw B. wan sani.*
but uh I FUT ask madame B. one thing
‘But, uh, I wish to ask Mrs. B. something.’ (49A:23)
- b. *Mi taki neen, mi wani tan, mi sa tan wroko.*
I say no I want stay I FUT stay work
‘I said no, I want to stay, I would stay and work.’ (49A:30)

The differences between *o* and *sa* become even clearer when we consider certain secondary uses of *sa* which refer to past events, as Voorhoeve (1957) pointed out. According to Wendelaar & Koefoed (1988:7), ‘*sa* adds to these sentences an element of doubt or indicates that unrealities are involved. This pure modal use is impossible for *o*’. My data contain only two examples of this use of *sa*, as follows.

- (20) a. A: *Neen, mi no sabi en.*
no I NEG know him
B: *Ma oom N. sa yere fu en, tog?*
but uncle N. FUT hear of him TAG
‘No, I don’t know him.
But uncle N. may have heard of him, right?’ (39A: 04)
- b. *Mi e kari a suma kon ... ma ala leisi a suma*
I IMP call the person come but all time the person

e poti en fesi go a oost. Dan mi taigi a suma,
 IMP put her face go the east then I tell the person
dan a suma sa piki mi wan sani... Ma di
 then the person FUT speak me one thing but when
mi miti B. meki trobi. B. e taki libi a sma
 I meet B. make trouble B. IMP talk leave the person
yungu.
 boy

‘I kept calling the girl, but every time the girl would turn her face away. Then I told her something, and the girl would tell me something... But when I met B. there was trouble. B. kept saying leave the girl alone, man.’ (16A: 33)

There are also several examples in my data where *o* marks future in relation not to S, but to some point in the past. The following are examples:

- (21) a. *Mi taigi en meki a sukuwan souvenir gi a man;*
 I tell him make he seek one souvenir for the man
a man taki a o sukuwan vaas gi en.
 the man say he FUT seek one vase for him
 ‘I told him to look for a souvenir for the guy: he said he would look for a vase for him.’ (8A: 19)
- b. *A ben taigi mi a o kon na fesisey baka.*
 he PAST tell me he FUT come LOC frontside back
 ‘He told me he would come to the front again.’ (8A: 24)

This suggests that *o* marks relative future, just as *ben* marks relative past, and that the tense system of SN in general is a relative tense system. In this respect SN is similar to GC, Jamaican Creole and other varieties of CEC, which also employ their future markers in this way (Winford 1993: 60).

7.3. Secondary uses of future *o*

As we saw earlier, in its primary uses, Future *o*, like future tenses crosslinguistically, conveys a sense of prediction by the speaker, as well as a sense of intention in cases where the speaker is the agent. There are also certain secondary uses of *o* which have parallels crosslinguistically. For instance, it often expresses a sense of predictability or characteristic behavior in both present and past contexts:

- (22) a. *Vanaf munde uit yu e nyan supu, heri heri,*
 as of Monday out you IMP eat soup whole whole
noem maar op. Sonde den suma o bori aleisi.
 name but on Sunday the-PL person FUT cook rice
Iya so a sani ben de.
 yes so the thing PAST COP
 'From Monday on, you would eat soup, heri heri [a dish of plantains, fish, etc.] and so on. On Sunday, people would cook rice. Yes, that's how it was.' (34B: 13)
- b. *Yu o wasi a krosi, wan heri tobo krosi nanga*
 you FUT wash the clothes one whole tub clothes with
wasi planga. Omeni a suma o gi yu?
 wash board how.much the person FUT give you
Vyftig cent. Ma yu ben de tevreden.
 fifty cent but you PAST COP content
 'You'd wash the clothes, a whole tub of clothes with a scrubbing board. How much would the lady give you? Fifty cents. But you were satisfied.' (49A: 20)
- c. *Den efu den suma fu disi sei e pley a Totnes*
 then if the-PL person of this side IMP play LOC Totnes
nanga den man dati, dan noiti den o kraak,
 with the-PL man DEM then never they FUT cheer
hori kraak fu disi sei wini.
 root for this side win
 'Then if the people from here are playing in Totnes against those people, then they will never root for the people from here to win.' (39A:44)

Similar to these examples is the use of *o* to express a routine procedure, as in the following instructions for a recipe.

- (23) *Nanga wan pikin gritigriti yu o griti a amandra. Yu*
 with one small grater you FUT grate the almond you
o piri en dan yu o poti en na watra. Yu musu
 FUT peel it then you FUT put it LOC water you must
moksi en, bita amandra nanga switi amandra.
 mix it bitter almond with sweet almond
 'You grate the almond with a small grater. You peel it, then you put it in water. You must mix it, bitter and sweet almond.' (1B: 28)

The latter secondary uses are rather rare. As we shall see, Imperfective *e* is much more commonly used to express these meanings.

7.4. Prospective *e go*

As I noted earlier, apart from the two future markers just discussed, SN also employs the construction *e + go* to convey the sense of an immediate or prospective future. Voorhoeve's (1957) analysis of *e go* as containing (progressive) *e* plus main verbal *go* is accurate, as is his conclusion that *e go* is not fully grammaticized as a tense category in SN. In fact, many uses of *e go* retain a quite literal sense of movement without any suggestion of futurity, as in the following examples:

- (24) a. *Ala mun yu e kisi yu pensioen ma ala mun*
 all month you IMP get your pension but all month
yu e go teki yu moni na bangi.
 you IMP go take your money LOC bank
 'Every month you get your pension, and every month you go
 and take your money at the bank.' (16B: 20)
- b. *Dan a so den ben e kon aksi mi krosi*
 then COP so they PAST IMP come ask me clothes
fu den pikin. Dan den e go nai tyari gi mi.
 of the-PL child then they IMP go sew carry give me
 'And so they would come and ask me for the children's clothes.
 And they would go and sew them and bring them to me.' (34A:
 17)
- c. *A bun dan, suma e go seti sani poti dape.*
 COP good then person IMP go set thing put there
Nownow grontapu kon frede.
 now world come afraid
 'Well, alright, people go and practise black magic there. To-
 day the world has become afraid.' (12A: 24)

In other cases, however, the sense of futurity is more in evidence.

- (25) a. *Efu yu wani wan man, fa yu e go teki wan man?*
 if you want one man how you IMP go take one man
 'If you want a man, how are you going to take a man?' (12B:
 11)

- b. *Te yu mama no de, kon mi taki a meki wan*
 when your mother NEG COP come I say she make one
switi sani, yu no e go fufuru en?
 sweet thing you NEG IMP go steal it
 ‘When your mom’s not there, let’s say she makes something
 sweet, won’t you steal it?’ (16A: 23)

The sense conveyed by *e go* in these cases is reminiscent of the English prospective future *be going to*. As Bybee & Dahl suggest with respect to the latter,

It is typical of *go* futures that they comprise a semantic element indicating ‘movement’, one indicating ‘toward a goal’ put into an imperfective or progressive aspect, yielding a literal meaning of ‘an agent is on a path toward a goal.’ (1989: 91)

Judging from the examples given above, the meaning of *e go* is still restricted to the sense of actual movement by an animate agent. Unlike English *be going to*, it has not apparently been grammaticized to the point where it can refer to future events involving non-animate entities or predictions like the following, where no literal movement occurs:

- (26) a. The stick is going to break.
 b. Susan’s going to be here tomorrow.

Possible evidence that *e go* is becoming more fully grammaticized comes from the following example, the only one in my data, where *e go* precedes a modal, and appears to convey simple futurity rather than the literal sense of movement. The example is somewhat unusual in that the speaker is quoting a song.

- (27) A. *Ayi, owma sabi a singi. Lena fu Mak.... Suma ben*
 yes granny know the song Lena of Mak... who PAST
singi..?
 sing
 B. *Lena fu Maka Olo, Lena yu e go kiri mi. Ohsani*
 Lena of Maka Hole Lena you IMP go kill me oh thing
dati, wansi wan bruku yu no e go bai. Sibi a
 that even one panty you NEG IMP go buy sweep the
oso yu no e go man, bori wan patu yu no
 house you NEG IMP go can cook one pot you NEG

e go man, ayi, taki mama, mi wani wan man.

IMP go can yes talk mama I want a man

- A. 'Yes, granny knows that song. 'Lena of Mak.. who sang that?'
- B. 'Lena of Maka Olo, Lena you're going to kill me. Oh, that thing, you won't even buy a panty, you won't be able to sweep the house, you won't be able to cook, yes, talkative mama, I want a man.' (12B: 11)

It is quite possible that the *e go* construction was the source of Future *o* via a process of grammaticization. It is common in such cases that both the original construction and its fully grammaticized counterpart continue to co-exist side by side. In the light of this, Voorhoeve's failure to distinguish the two is perhaps understandable.

8. Aspect in Sranan

As is well known, aspect characterizes the internal structure of a situation, or, in Comrie's (1976:4) terms, "the internal temporal constituency of the situation". Chung & Timberlake suggest more specifically that "aspect characterizes the relationship of a predicate to the time interval over which it occurs" (1985:213). In general, situations may be characterized in terms of aspectual notions such as dynamicity (e.g., progressive/non-progressive), closure (e.g., perfective/imperfective), iterativity and habituality, durativity and so on. Some of these notions (e.g. dynamicity and closure) are also relevant to inherent aspectual properties of the predicates themselves. Thus particular verbs may describe situations that are inherently either static or dynamic, durative or non-durative, telic or non-telic and so on. It is therefore of crucial importance to keep the notion of an aspectual category distinct from the "aspectual potential" of verbal lexemes. As was pointed out before, failure to do so is "a constant source of confusion in the study of TMA categories" (Dahl 1985).

As noted earlier, the discussion of aspect in the creole literature is a good example of the terminological and methodological inconsistency in the field. Bickerton's 'prototypical' creole TMA system included only one aspectual category, to which he applied the label 'non-punctual'. This label was intended to subsume meanings such as 'progressive-durative plus habitual-iterative' (1981:58). His examples include aspectual *a* in GC and *e* in SN. Most creolists have adopted this label, sometimes applying it to categories

which do not express the same range of meanings. Moreover, they have tended to overlook other aspectual categories which are just as integral a part of creole tense/aspect systems.

Bickerton's analysis of the 'punctual/non-punctual' distinction has been found wanting in several respects. First, as scholars like Andersen (1990:63) have pointed out, labels like '(non-)punctual' apply more properly to the inherent aspectual properties of predicates as distinct from aspectual categories. Bickerton's analysis fails to make this distinction, and treats '(non-)punctuality' both as a property of predicates and a manifestation of a TMA category. This leads to an inaccurate interpretation of the interaction between the 'non-punctual' category itself and stative predicates in GC, as when he claims that non-punctual *a* is compatible only with non-statives, since a stative is by definition non-punctual. In fact, as Gibson (1982) Jaganauth (1987) and Winford (1993) have shown, auxiliary *a* can occur with statives in GC, but then its meaning is 'habitual', whereas it may be interpreted as either progressive or habitual with non-statives. Since Bickerton's label 'non-punctual' was intended to apply to other creoles, including SN, the inaccuracies in his analysis extend to these creoles as well. I am in agreement with Andersen (1990) and others that the label 'Imperfective' is more appropriate than 'non-punctual' for the relevant category in GC, SN and other creoles. I will present evidence below that SN *e* is a classic example of an Imperfective marker.

Apart from this, the attention devoted to Bickerton's 'non-punctual' category resulted in the neglect of other categories that are just as vital to creole aspectual systems. A case in point is the Completive category, expressed by *don* (< *done*) in all varieties of CEC, by *feni* (< *finish*) in other Atlantic English creoles, by *kaba* (< Portuguese/Spanish *acabar* 'finish') in Sranan and other Surinamese creoles, as well as in Iberian-based creoles (Stolz 1987), and by reflexes of *finir* 'finish' in French-based creoles. Bickerton excluded this category from his prototypical TMA system. But Boretzky has argued that it is quite common among creoles of different lexical affiliation throughout the world, and hence should be included as part of the prototype. I agree, and shall argue that Completive *kaba* (normally pronounced *kba*) is an important part of the SN aspectual system, though it appears to be undergoing reanalysis from verb to adverb. In both its semantics and use, it shows clear parallels to Completive grams in other creoles.

8.1. Imperfective *e* in Sranan

Previous analyses of auxiliary *e* in SN have noted that it expresses a range of meanings including ‘durative, iterative and continuous’ (Voorhoeve 1957: 376; Seuren 1981:1052). This range corresponds closely to that typically associated with the category-type IMPERFECTIVE cross-linguistically. In general, this category presents a situation as unbounded. More specifically, as Bybee et al. (1994:125) note, “an imperfective situation may be one viewed as in progress at a particular reference point, either in the past or present, or one viewed as characteristic of a period of time that includes the reference time, that is, a habitual situation”.

Pace Seuren (1981:1047), both stative and non-stative predicates in SN are regularly marked by *e*, though the interaction between *Aktionsart* and the category leads to somewhat different interpretations in each case. This is as good a point as any to discuss the classification of predicates in SN according to stativity, since it is crucial to the operation of aspect in particular. As Bybee et al. (1994:55) point out,

a stative predicate describes an unchanging situation which will continue unless something happens to change it, such as *know*, *want*, *be tall*, *be ripe*, *be located*. A dynamic predicate typically describes a situation which involves some sort of change (Comrie 1976:48-50), such as *write*, *walk*, *sneeze*, *ripen*, *drop*.

Seuren’s (1981) distinction between what he calls ‘A-verbs’ and ‘E-verbs’ in SN corresponds roughly to the state/non-state distinction. The former class includes verbs such as *abi* ‘have’, *lobi* ‘like, love’, *sabi* ‘know’, *wani* ‘want’, etc, as well as auxiliary verbs like *kan* ‘can’, *musu* ‘must’, *sa* ‘may’ etc. The latter category is much larger and more open, and includes verbs such as *nyan* ‘eat’, *waka* ‘walk’, *lon* ‘run’, *skrifɪ* ‘write’, etc. Seuren’s classification is by and large correct, though his criterion for distinguishing the two classes, namely that E-verbs require *e* in some cases while the use of *e* is not normally found with A-verbs (1981:1047-48), is inaccurate.

To illustrate the use of *e*, we may note first of all that when combined with dynamic predicates it often conveys the sense of an activity or event in progress, ie, ‘progressive’ meaning. The following examples illustrate:

- (28) a. *Nownow yu e teki en kba nownow?*
 now you IMP take it already now
 ‘Are you already taping right now?’ (Tape 39A-1)

- b. *Ma den suma san e libi dya a abra sei*
 but the-PL person REL IMP live here LOC other.side
na den M tog? Na den e libi dape?
 COP the-PL M TAG COP they IMP live there
 ‘But the people who live across the street are the M’s, right? Is it they who live there?’ (Tape 39A-31)
- c. *Ma dan kon unu libi a tori dati yere, want a*
 but then come we leave the story DEM hear for the
sani disi na gro skin, want ala mi skin e gro
 thing DEM COP grow skin for all my skin IMP grow
kba.
 already
 ‘But let’s forget that story, right, because that stuff causes goose bumps, for my whole body is full of goose bumps already.’ (Tape 16A-22)
- d. *Mi e sidon dya, mi e denki fu beroof wan bangi*
 I IMP sit here I IMP think to rob one bank
now de.
 now there
 ‘I’m sitting here thinking to rob a bank right now.’ (Tape 16A-22)

Imperfective may be used “to describe activities that are not actually in progress at reference time, but that are characteristic of a certain time frame which includes the reference time” (compare ‘extended’ uses of the Progressive in English) (Bybee & Dahl 1989:82).

- (29) a. *Wan tu fu den pikin fu owma e wroko gron*
 one two of the-PL child of granny IMP work ground
now ooktu?
 now also
 ‘Are some of granny’s children also cultivating the land now?’
 (Tape 34A-11)
- b. *Of a man e fokop a kondre tu. Luki dya,*
 if the man IMP fuck-up the country too look here
pikin pikin negre, yu sabi omeni suma e
 small small child you know how.many person IMP

pina ini Sranan? Den man e pina,
 suffer in Suriname the-PL man IMP suffer
yungu. A sari yungu. Traman e abi angri.
 man COP sad man. other IMP have hunger
Traman no e nyan srefisrefi.
 other NEG IMP eat selfself

‘The guy’s really fucking up the country. Look here, small children, do you know how many people are suffering in Suriname? Those people are suffering, man. It’s sad, man. Others are going hungry. Others aren’t eating at all.’ (Tape A-26)

- c. R. *Na yu eigi wagi disi?*

COP your own car this

(‘Is this your own car?’) (Tape 39A-29)

G. *Neen. Na a heer disi e ry mi. Mi musu pai*
 no COP the man DEM IMP drive me I must pay
en.

him

‘No, it’s this gentleman that’s driving me around. I have to pay him.’

Just as commonly, *e* is used to describe repeated or habitual situations with both statives and non-statives:

- (30) a. *En te den man dati e kon tapu a dan,*
 and when the-PL man DEM IMP come top the street
dan den e law. Den man echt! Altijd unu echt! No
 then they IMP mad the-PL man real. always we real no
wan tra band e sidon na wi sey.
 one other band IMP sit LOC our side

‘And when those guys come out on the street, they go crazy. Those guys are it! Always we’re it! No other band can match us.’ (Tape 16A-17)

- b. *Want son leisi yu abi den suma san, uh,*
 for some time you have the-PL person REL uh
den e afrontu den tongo, tog.
 they IMP turn.against their language TAG

Den no e wani taki Sranan Tongo moro.

they NEG IMP want talk Suriname tongue more

'For sometimes you have people who, uh, turn their backs on their language, right. They don't want to speak Sranan Tongo anymore.' (Tape 34B-27).

- c. *En te yu si wan moi film a tv, yu e wan go luku. Yu e prakseri a schrift, yu e bonk a schrift go na wan sey. Yu no e abi ten.*

and when you see one nice film LOC tv you IMP want go look you IMP study the notebook you IMP throw the notebook go LOC one side you NEG IMP have time

'And when you see a nice movie on tv, you want to go watch it. You consider the notebook, you throw the notebook aside. You don't have time.' (Tape 16B-27)

Closely related to the above examples are those in which *e* is used to describe generic truths, and those where it describes routine or prescribed activity (as in directions for a recipe):

- (31) a. *Fowru e singi, dagu e bari.*

bird IMP sing dog IMP bark

'Birds sing, dogs bark.' (E-13)

- b. *Want a kokori yu e fon en. Dan te yu fon en yu e zeef en. Dan te yu zeef en dan yu e go poti yu watra a faya ini wan patu. Dan a watra e kuku....*

because the kokori you IMP beat it then when you beat it you IMP strain it than when you strain it then you IMP go put your water LOC fire in one pot then the water IMP cook

'Because you mash the kokori [kind of oatmeal]. Then when you've mashed it you strain it. Then when you strain it, you put your water to heat in a pot. Then the water boils....' (12B: 21)

To describe generic situations of this kind, as noted earlier, Future *o* often alternates with *e*, as in the following example.

- (32) *Dan yu e klop en dan te yu klop en dan yu e*
 then you IMP stir it then when you stir it then you IMP
si alamala e kon bruin. Dan yu o poti pikin
 see everything IMP come brown then you FUT put little
watra gi en dan yu o opo a tonton ondro fu a
 water for it then you FUT lift the tonton under for it
no bron.

NEG burn

‘Then you stir it and when you have stirred it then you see every-
 thing turn brown. Then you put a little water in it and you lift the
 tonton [mashed plantains] from below so it doesn’t burn.’ (12B:
 22)

It is also possible for *e* to mark modal auxiliaries, as in the following ex-
 amples.

- (33) a. *Ma tori e de fu taki, ma yu no e man*
 but story IMP COP for talk but you NEG IMP can
onthou ala den tori so moro.
 remember all the-PL story so more
 ‘But there are stories to tell, but you can’t remember the stories
 anymore.’ (39A: 2)
- b. *Oh, yu e meki bestselling, dus, dati suma e kan*
 oh you IMP make orders so that person IMP can
kon bestel a yu of yu e meki gewoonfu
 come order LOC you or you IMP make only for
yu oso?
 your house
 ‘Oh, do you take orders, so that people can order from you, or
 do you make [cakes] just for yourself?’ (39A: 43)
- c. *Dus dan mi e taigi mi pikin, mi meki muyti*
 so then I IMP tell my children I make effort
nanga den. Dus suma pikin e musu meki muyti.
 with them so person small IMP must make effort
 ‘So I would tell my children that I had done my best with them.
 So the children must do their best too.’ (49B: 9)

The use of *e* to express habitual meaning is quite common with stative
 predicates, as the examples above demonstrate, but this has been generally

(34) a. *Safrisafri mi e* *sabi den pasi kba.*
slowly I am.beginning.to know the roads already
b. *Alen e nati* *yu.*
rain makes.wet you
c. *Pe* *Srananman* *e* *de, yu abi prisiri.*
where Surinam.people ever are you have fun
d. *Yu no e* *sabi a* *wet so bun.*
you not really know the law so well
(Voorhoeve 1957: 378)

I would suggest, first of all, that no form of lexical derivation is involved in any of these cases. Sentences like b) can be better explained by noting that property items like *nati* 'wet', *faya* 'hot' etc. can function as both adjectives and verbs in Sranan, as in other creoles. In their verbal function, they are processual, and are compatible with *e* in either its progressive or habitual sense (Winford 1997). There is no need to posit a derivational rule triggered by the appearance of *e* that would convert 'adjectives' to 'verbs' in these cases.¹⁶ Secondly, sentences like c), involving the habitual interpretation of *e* with statives, are quite common in Sranan, as in other creoles, and require no special derivational process to explain them. They simply illustrate the typical use of an Imperfective to convey habituality, as illustrated earlier. Similarly, the use of *e* in examples like d) is in keeping with its function as an Imperfective marker, and simply conveys the sense of habituality or a continuing situation. The sense of 'intensive' which Seuren reads into this use is more a matter of the implicatures he appears to find in the sentence, than of the true

meaning of *e*. Examples like a), however, are rather less usual, and Seuren is essentially correct in noting that the stative predicates in these cases seem to be interpreted in somewhat dynamic terms. For instance, in the following examples from my data, predicates like *kowru* and *frede* describe situations conceived of by the speaker as involving a process or change:

- (35) a. *Yu e kanti en ini a koba, a papa, yu kanti*
 you IMP pour it in the bowl, the porridge you pour
en go ini a koba, dan a e kowru dan a e puru
 it go in the bowl then it IMP cold then it IMP pull
na vorm fu a koba.
 the form of the bowl
 'You pour it into the bowl, the porridge, you pour it in the bowl, and it gets cold and then takes the shape of the bowl.' (12B: 25)
- b. *Wan leysi mi si wan weri ede sani ma mi no e*
 one time I see one weary head thing but I NEG IMP
frede, yere. Ma yu e firi dati yu ati e
 be.afraid hear but you IMP feel that your heart IMP
naki pikinso.
 knock a little
 'Once I saw a terrible thing, but I wasn't scared, right. But you feel your heart beating a bit faster.' (8A: 10)

Cases like these are analogous to examples found in English and other languages, where states can be presented in dynamic terms and marked by progressive, as in the following:

- (36) a. You're being difficult.
 b. I'm understanding math better every day.

Like other languages, SN allows for a certain degree of flexibility in the behavior of its predicates, particularly those that describe properties and changes of state (Winford 1997).

Finally, it must be emphasized that Imperfective *e* represents a purely aspectual category, quite neutral with respect to time reference. In the examples presented so far, the reference point is S, and hence *e* is interpreted as having present time reference. But it is equally common to find *e* used in contexts where the time reference has been established as some point in the past, for instance by Past *ben* or by adverbials. The following examples illustrate:

- (37) a. *Ma dan laast laast mi e luku tapu tv, wan*
 but then lately lately I IMP look top tv one
aardbeving, boi. Omeni suma dede boi.
 earthquake boy how.many person die boy
 'But then recently I was watching tv..an earthquake, boy. How
 many people died, boy.' (Tape 16A: 21)
- b. *Ma di mi ben e gro kon, mi no ben mag*
 but when I PAST IMP grow come I NEG PAST may
taigi wan bigi suma wan wortu. Uh? Efu mi seni a
 tell one big person one word uh if I send the
pikin a no go, en mama yere, a e fon en. Taki
 child he NEG go his mother hear she IMP beat him say
san.ede meki te owma seni yu, yu no go?
 why make when granny send you you NEG go
Direct a e priti en skin gi en.
 immediately she IMP cut her skin for her
 'But when I was growing up, I wasn't allowed to say a word to
 an adult. Uh? If I sent a child [on an errand] and s/he didn't go,
 and his/her mother heard this, she would spank the child.
 [She'd] say why didn't you go when granny sent you? Right
 then she'd cut her skin for her.' (34A: 30)
- c. *Faya no ben de. No wan gado faya. Kronto bon*
 fire NEG PAST COP NEG one god fire Coconut tree
ben furu a strati. Faya no de. A so te neti
 PAST full LOC street fire NEG COP COP so when night
yu e waka. Efu yu no abi flashlait efu
 you IMP walk if you NEG have flashlight or
munkenki, yu no e si. Yu e djam srefi nanga
 moonlight you NEG IMP see you IMP jam even with
suma.
 person
 'There were no lights. Not a single light. There were lots of
 coconut trees along the street. There were no lights. That's how
 it was when you walked at night. If you didn't have a flashlight
 or moonlight, you couldn't see. You'd even bump into people.'
 (49A: 12)

Similar to the above examples are cases involving the use of *e* in embedded clauses, where *e* picks up past time reference from the matrix verb, as in the following examples.

- (38) a. *Mi pikin no ben e taigi wan ouders taki a*
 my child NEG PAST IMP tell one elder COMP s/he
no e go.
 NEG IMP go
 'My children would never tell one of their parents that they weren't going.' (34A: 31)
- b. *Want di wi e pasa a oso ben tapu.*
 because when we IMP pass the house PAST lock
 'Because when we were passing by the house was locked.'
 (39A: 10)
- c. *Dan a man taki wan neti a sidon nanga a kil*
 then the man say one night he sit with the guy
ma dan a kil ben de bezig e bori.
 but then the guy PAST COP busy IMP cook
 'Then the man said one night he sat with the guy, and the guy was busy cooking.' (8A: 3)

Some scholars, e.g., Mufwene (1984), have taken examples such as these in other Caribbean English creoles to support the view that these creoles have a 'relative tense system.' I have already shown that Past and Future tenses in Sranan are indeed relative tenses. But it may be more accurate to say that the verbs in the embedded clauses in these cases display dependent time reference, similar to what occurs in other cases of embedding in CEC, for instance in complements to perception verbs, desideratives and so on (Winford 1993).¹⁷ The same applies to SN *e*, and is in keeping with the view expressed earlier that the Imperfective marker is neutral with respect to time reference, or, perhaps more accurately, expresses time reference simultaneous to that established in the discourse context.

Finally, we may note that, in its progressive function, *e* sometimes conveys the sense of an immediate future, like progressive grams crosslinguistically:

- (39) a. *Dan mi teki wan owru taki yungu efu yu no go*
 then I take one machete say man if you NEG go

- a strati mi e kapu yu kiri yungu.*
 LOC street I IMP chop you kill man
 ‘Then I took up a machete and said “man, if you don’t get on the street I’ll chop you to death man”.’ (39A: 18)
- b. *Mi taki a fu dati yu e krei? Mi taki dan mi*
 I say COP for that you IMP cry I say then I
e puru a nyan na yu anu dan mi e gi
 IMP pull the food LOC your hand then I IMP give
den brada.
 the-PL brother
 ‘I said ‘is that why you’re crying?’ I said then I’ll take the food away from you and give it to your brothers.’ (34B: 1)

This represents a secondary use of *e*, and is rather restricted in my data. It would appear that some creoles such as Haitian creole (Spears 1990), employ their progressive markers to express futurity much more frequently than SN does. This may be an indication of change in these creoles in which a previously secondary function of the progressive marker has become primary, but further data are needed to confirm this conclusion.

8.2. *Completive kaba*

Sranan, like other creoles, has a Completive marker (realized as *kaba* or *kba*) which is formally similar to a main verb (also realized as *kaba* or *kba*) which means ‘finish.’ Sranan and the other Surinamese creoles may have adopted *kaba* from Portuguese *acabar* ‘finish’ or it may be that *kaba* was borrowed into the other creoles either from a Portuguese pidgin or from Saramaccan, in whose formation Portuguese lexical input played a significant role. In varieties of CEC, the completive marker and the main verb meaning ‘finish’ are expressed by *don* (< English *done*).

I shall argue here that the Completive marker shows evidence of being grammaticized into an adverb in SN, but almost certainly originated as a serial verb meaning ‘finish.’ It now conveys the sense of ‘already.’ It typically occurs in VP-final position, reflecting its serial verb origins, but can also appear at the end of other types of phrase, a fact which I interpret to be part of its development into a true adverbial. The fact that VP-final *kaba* seems to be phonologically reduced more often than its main-verbal relative is a further reflection of this grammaticization. This conclusion is basically in agreement

with that of Seuren (1981: 1074, fn 5) who notes that “*kba* (from Portuguese/Spanish *acabar* ‘finish’) can be used in two ways; as a verb, followed by an infinitive, meaning ‘finish’: *mi kba nyam* ‘I have finished eating’, or as an adverb meaning ‘already’, in which case it stands at the end of the sentence: *mi nyam kba* ‘I have eaten already’. It is the obvious counterpart of *don* (from English *done*) in the neighbouring Guyanese creole”. Voorhoeve (1957) makes no mention of *kaba* in his account of the TMA system of SN. Nickel & Wilner (1984) mention it, but fail to distinguish clearly between Completive *kaba* and main verbal *kaba*. They note that “the adverbial *kba* ‘already’ or ‘finished’, which normally appears clause-finally, may be fronted to the auxiliary position” (1984: 43). I agree with Seuren that the two are quite distinct.

Seuren does not include either preverbal *kaba* or VP-final *kaba* as part of the TMA system of SN. His reason seems to be that neither is strictly speaking a member of the AUX category which he posits as the locus of TMA auxiliaries for that creole. However, it is important to include these items in the present discussion, first, because they do in fact play an important role in the aspectual system of SN, and secondly, because they have parallels in other creoles and are therefore important to our understanding of cross-creole similarities and differences in this area.¹⁸ We may note first of all that the main verb *kaba* can appear either as a transitive verb with an NP object, as a main verb with a VP complement, or as an intransitive verb. Note that in these functions *kaba* may be preceded by TMA auxiliaries. The following examples illustrate:

- (40) a. *Te mi mama komoto a gron kon, dan a*
 when my mother come LOC ground come then the
 nyan kaba.
 food finish
 ‘By the time my mother came back from the garden, the food
 was finished’ (34A: 5)
- b. *Pas te unu kaba nanga skoro dan wi o meki*
 only when we finish with school then we FUT make
 pikin nanga den sani dati.
 child and the-PL thing DEM
 ‘Only when we finish with school, then we’ll have kids and
 that sort of thing.’ (16B: 12).

- c. *Dus... a man ben kaba rij, a man ben kaba*
 so the man PAST finish ride, the man PAST finish
rij helemaal Ma dan wan laatst pikin aksi a man
 ride completely but then one last girl ask the man
fu a man rij a laatst rij.
 for the man ride a last ride
 'So.. the man had finished driving, he had finished driving
 completely. But then this last girl asked him to make one last
 trip.' (16B: 35)
- d. *Te yu e drai baka so, yu e taki a bigi suma*
 when you IMP turn back so you IMP tell the big person
odi pasa tin tron, want a no musu taigi yu
 howdy pass ten time because s/he NEG must tell your
mama taki a ben si yu nanga a jongen. So yu
 mother say s/he PAST see you with the boy so you
si, yu no e kaba taki en odi nownow.
 see you NEG IMP finish tell her howdy now
 'When you turn around, you say hello to the adult many times,
 for she mustn't tell your mom that she saw you with the boy. So
 you never stop saying hello to her.' (49A: 54)

Preverbal *kaba* is always a main verb, with the sense of 'finish'. By contrast, preverbal *don* in CEC may be either main-verbal ('finish') or it may be an auxiliary marking Completive aspect, expressing the sense of 'already' and functioning in ways quite similar to a type of PERFECT. The latter type of meaning can be conveyed in SN only by VP-final *kaba*. Instances of the latter are quite common in my data, occurring with both non-statives and statives (including progressive and adjectival predicates). With non-statives, it conveys the sense of a past event that leads to some result with implications for the current situation. in this use, *kaba* conveys a meaning similar to that of a perfect of result. The following examples illustrate:

- (41) a. *Want na tu leisi mi nanga a man meki afspraak*
 because COP two time I and the man make appointment
kaba, a man no kon.
 already the man NEG come
 'Because it's twice that me and the guy made appointments
 already, and he never came.' (8A: 22)

- b. *A alen disi kan stop now. Yongu, a kon tumsi furu*
 The rain this can stop now Man it come too full
kaba, yere.
 already hear
 'This rain can stop now. Man, it has already rained more than enough.' (8A: 29)
- c. *Te prani aleisi ten, mi ben e gwe go prani aleisi*
 when plant rice time I PASTIMP go go plant rice
want Gado ben gi mi pikin kaba.
 because God PAST give me child already
 'When it was time to plant rice, I went to plant rice, for God had already given me children.' (49A: 8)
- d. *Dus mi ben go a wan tu suma kaba. Dan mi*
 so I PAST go LOC one two person already then I
prakseri kon mi kon na oom N.
 think come I come LOC uncle N
 'So I'd already gone to one or two people. Then I thought, let me come to Uncle N.' (39A: 30)

In these examples, *kaba* has a similar meaning and function to VP-final *don* in CEC, which marks some event as having been completed at S. I follow here the practice of creolists in referring to *don* and *kaba* as markers of Completive aspect, with the added stipulation that Completive is a sub-type or close relative of PERFECT.¹⁹ Like typical perfects, it signals that "the situation occurs prior to reference time and is relevant to the situation at reference time" (Bybee et al. 1994: 54). This sense is compatible with the sense of adverbials like 'already' which often accompany perfects. Indeed, as suggested earlier, it appears that *kaba* itself is on the way to being grammaticized into just such a relational adverb. To elaborate, it seems that VP-final *kaba*, like VP-final *don*, originated as a serial verb meaning 'finish', then developed into a type of perfect marker. This path of development has been widely attested in TMA systems crosslinguistically (Bybee et al. 1994: 70; Bybee & Dahl 1989: 68). The further evolution of such grams into adverbials is not as well documented.

The uses of VP-final *kaba* with non-statives illustrated above resemble the prototypical use of PERFECT to describe a past action with current relevance. In addition, however, *kba* is used with stative verbs (including progressives, adjectivals and passives) to convey the sense that the state is

already in existence and has been for some time, with some implication for the present situation. This type of meaning is quite similar to that of a resultative (Bybee et al. 1994: 74). The slight difference in interpretation of *kaba* with statives vs. non-statives can be accounted for in terms of differences in *Aktionsart*, as we have seen for other TMA categories. The following examples illustrate:

- (42) a. *Ma yu sabi fa den suma fu Flora de kba.*
 but you know how the-PL person of Flora COP already
Den lobi ogri.
 they love ugliness
 'But you know what Flora people are like already. They love wickedness.' (16A: 24)
- b. *Te wan mama abi pikin kba, a pikin musu nyan.*
 when one mother have child already the child must eat
 'When a mother has children already, the children must eat.'
 (49A: 9)
- c. *Den suma no abi deki ati tog, noso*
 the-PL person NEG have thick heart TAG otherwise
den sani disi ben post o langa kba.
 the-PL thing this PAST mail how long already
 'They don't have courage, right, otherwise these things would have been mailed already.' (8A: 20)
- d. *Ohma yu musu du wan sport yere want yu si*
 oh but you must do one sport hear because you see
fa yu fini fini kba; yu musu meki yu kon
 how you fine fine already you must make you come
diki diki pikinso.
 thick thick a.little
 'Oh, but you must practice a sport right, because you see how you're skinny already, you must let yourself get a little big-gish.' (16A: 16)

Examples like these, which are quite common in my data, are evidence that *kaba* has reached an advanced stage of semantic development in SN, and has been generalized to all types of predicate. In this respect, it is comparable to auxiliary and VP-final *don* in GC, and to auxiliary *don* in other varieties of CEC. One exception is Jamaican Creole auxiliary and VP-final *don*, which

still retains a strong terminative sense and is compatible only with non-statives and change of state predicates (Winford 1993: 48). This suggests that Jamaican Creole *don* is in an early stage of development when compared with its counterparts in other varieties of CEC.

I suggested earlier that SN *kaba* has advanced even further on the path of grammaticization, to the point where it appears to modify adverbial phrases of time. The following examples illustrate:

- (43) a. *A sani disi a wan sani fu dertien, veertien*
 thething this COP one thing of thirteen fourteen
yari kaba.
 year already
 'This is something that happened thirteen or fourteen years ago.' (39A: 18)
- b. *Kon unu taki, dus, den famiri no libi, no e libi*
 come we say so the-PL family NEG live NEG IMP live
dyaso moro; jaren kaba den no e libi dyaso.
 here more years already they NEG IMP live here
 'Let's say, so, the family hasn't lived, don't live here anymore;
 it's years now they haven't been living here.' (34B: 54)

In these cases, *kaba* appears to function like an adverbial.

9. Conclusion

I have argued that the so-called "prototypical" TMA system which Bickerton proposed, and which creolists have used as a measure of similarities and differences among creoles, is flawed in many respects. His focus on just three core categories and the unmarked verb neglects several other categories which are integral parts of radical creole TMS systems, such as the Completive/Perfect, the Uncertain vs. Certain Future and several modal categories of Sranan (see Winford to appear). Even the analysis he proposes for the syntax of radical creole verb complexes, his well-known T-M-A ordering, is in need of revision. While Sranan allows such ordering, it permits other combinations as well. Moreover, other Caribbean English creoles (including GC, one of Bickerton's radical creoles) display quite different inventories and combinatory possibilities of TMA auxiliaries.

This is not to deny, of course, that there are striking similarities across creole TMA systems. The task of explaining such similarities and differences has only just begun, and it is crucial that researchers achieve common ground in terminology and methodology if our understanding of creole verb complexes (and their origins) is to improve. Based on available descriptions of creole TMA systems, it would appear that Sranan shares a number of core TMA categories with several other creoles, including those of English, French and Spanish lexicon, both in the Atlantic and Indian oceans. These categories include:

- Two tense categories: a relative Past and a predictive Future.
- The unmarked verb, conveying perfective aspect.
- Two aspectual categories: Imperfective and Completive-Perfect.
- Two primary modals: a modal of obligation equivalent to “must” and a modal of possibility equivalent to “can.”

However, these shared categories are not always identical in their range of meanings and uses across the creoles in question. Moreover, these very creoles differ in the inventory of other TMA categories they have, and in the syntactic properties of the TMA auxiliaries themselves. In addition, of course, there are creoles whose TMA systems depart even further from the core system referred to above.

Thus the growing literature on this area of creole grammar has made it abundantly clear that no creole conforms exactly to a single “prototypical” system. If the notion of a prototypical creole TMA system has any validity or value at all, it must be in the sense that many creoles share a certain core of TMA categories. This raises a number of questions which future research must address.

- How many creoles in fact share this common core? How many do not?
- What is the explanation for the common core?
- What categories other than the core categories does each creole have?
- How do we explain these additions to or departures from the common core?
- To what extent do the syntactic properties of TMA auxiliaries match across creoles, and why?

The answers to these questions will require much more research on the contemporary structure as well as the sources of creole TMA systems. Success

in achieving our goals will depend in the first place on the collection of fuller and more adequate data drawn from spontaneous speech, as well as careful elicitation in creole communities. Just as crucially, there is need for creolists to abandon the Bickertonian framework which has become the dominant paradigm for the investigation of creole TMA systems, in favor of more “empirically responsible” (Andersen 1990:89) analyses, which draw on established terminology and practice in the semantic and typological literature on TMA systems crosslinguistically.

Notes

- * The research on which this paper is based was supported by NSF Grant #SBR-930635, for which I wish to express my gratitude. I am also very grateful to the many informants who gave so freely of their time, to my principal fieldworker, Gladys Waterberg, without whom it would have been impossible to collect the data, to my research associate, Marinus Stephan, for the great part he played both in the fieldwork and transcription of the data, and to John Wilner and his associates at the Summer Institute for Linguistics in Paramaribo. Thanks also to Oesten Dahl, Mikael Parkvall and John McWhorter for their valuable comments on an earlier draft of this paper.
- 1. Dahl (1993: 251-52), reviewing Singler (1990) was struck by “the extent to which — due to the impact of Bickerton’s work — the study of creole TMA systems has become an autonomous tradition, with its own terminology and conceptual apparatus, with an ensuing relatively restricted influence of non-creolist TMA studies.”
- 2. It is important to emphasize that the data were collected from native speakers, since many Surinamese learn and use Sranan as a second language. Moreover, native speakers tend to be concentrated in working class communities of persons of African descent.
- 3. Indeed, much of the speech recorded, especially that of younger informants, is casual to the point that it contains slang and other highly idiomatic language that some second-language speakers of Sranan find difficult to understand.
- 4. I would like to express my sincere thanks to John Wilner, Norbert Rennert and their associates at SIL, Paramaribo, for all their valuable assistance. I’m especially grateful to Ronald Pinas and Marinus Stephan for their major contribution to the translations. Finally, my thanks go to Hertoch Linger, Margie McBean and several others for assistance with the elicitations and other tasks.
- 5. There is also some resemblance between Bickerton’s prototype and Taylor’s (1963) outline of Caribbean creole TMA systems.
- 6. Dahl (1985:74) argues, however, that the “totality” view “is not equally adequate for all language specific categories that it has been applied to” and suggests that the notion of boundedness is a central part of the Perfective category in languages like Russian. However, as Bybee & Dahl (1989:84) suggest, these two positions are not incompatible,

- since perfectivity is “a set of related concepts rather than one single notion,” and perfective categories “share a common focus, that is, the prototypical cases are the same.”
7. It is often difficult to distinguish between Dutch items which represent lexical code-switches (sometimes referred to as “nonce borrowings”) and those which are well integrated into the SN lexicon and treated like native words (eg. *moi* ‘nice, sweet, pretty’).
 8. The use of unmarked verbs in the *if*-clause of conditionals is also found in Romance-based creoles like Papiamentu. It has been claimed (e.g., by Andersen 1990 and Maurer 1993) that zero marking in such cases represents a category of ‘subjunctive’ mood. In my view, this interpretation is inaccurate, and reflects a tendency to interpret creole grammar in terms of the grammar of European languages like Spanish, where of course a subjunctive mood is used in similar cases.
 9. For instance, Bybee et al. (1994:54) use the term ‘anterior’ to refer to what have traditionally been called ‘perfects.’
 10. A quantitative study of past marking in SN narrative texts by Sankoff (1990:306) found that unmarked statives with past reference in fact outnumbered statives marked by *ben*, which is contrary to Bickerton’s claim that all past reference statives must be marked by *ben*.
 11. I am very grateful to Östen Dahl for pointing this out to me, and identifying several problematic cases in the data presented in an earlier draft of this paper.
 12. *Bo* derives from *ben* + Future *o*.
 13. Some quantitative studies of past marking in Sranan (Sankoff 1990) and other Caribbean creoles (Tagliamonte & Poplack 1988) have attempted to quantify uses of *ben* vs. unmarked verbs using Bickerton’s guidelines which state that *ben* marks past with statives and past before past with non-statives, while unmarked statives and non-statives have present and simple past reference respectively. It turns out however that these predictions are not borne out by my data. So far, no quantitative study has been attempted using the analysis offered here and in other studies such as Pollard (1989) and Wilner (1993). Previous studies did not have the benefit of these insights and hence had no clear basis on which to predict occurrences of *ben*. As Sankoff (1990:309) notes, “the main problem here is ... finding contexts where *ben* might have occurred but did not.” She acknowledges that “it was not possible to define the set of potential environments [for *ben*] explicitly” (ibid.)
 14. In fact, there is no known language that subsumes all such meanings under a single category. Bybee et al (1994:236fn) offer an informative critique of analyses which propose a realis/irrealis distinction as basic to the verb complex of certain Australian and New Guinean languages. They conclude that “realis/irrealis is rarely realized in a language as a binary morphological distinction. It appears to be more common to have multiple markers in both domains” (1994:238).
 15. By “(non-)realizable future” they appear to mean future events that are (not) subject to the control of the speaker. I am grateful to Marinus Stephan for translating this article from Dutch into English for me.

16. This is not to say, however, that some form of lexical derivational rule might not be appropriate to show the relationship between the adjectival and verbal functions of property items.
17. Andersen (1990:73) has noted a similar use of Imperfective *ta* in embedded clauses in Papiamentu, and concluded that its time reference must be inferred from the total context.
18. Stolz (1987: 303) agrees that what he refers to as “phrase- or sentence-final *kaba*” in SN and other creoles is an important part of these creoles’ TMA systems. However, he offers a rather different account of the relationship between pre-verbal and VP-final *kaba*, arguing that *kaba* is basically (underlyingly) an auxiliary, and that “there is still a certain liberty for *kaba* to occur in various syntactic positions without changing either its semantics or its grammatical status.” I do not think that this analysis is true for SN, or for the other creoles mentioned by Stolz.
19. Bybee et al. (1994: 54) define “Completive” in a somewhat different way, as “to do something thoroughly and to completion”, distinguishing it from what they call “anterior” (their label for what are usually called perfects) which are relational in function. In their approach, *don* and *kaba* would probably be treated as a type of Anterior.

References

- Anderson, L.B. 1982. “The ‘Perfect’ as a universal and as a language-specific category”. In *Tense-aspect: Between semantics and pragmatics*, ed. by Paul J. Hopper, 227-64. Amsterdam: John Benjamins.
- Andersen, Roger W. 1990. “Papiamentu tense-aspect, with special attention to discourse”. In Singler, 59-96.
- Bickerton, D. 1974. “Creolization, linguistic universals, natural semantax and the brain”. *Working Papers in Linguistics* 6: 125-41. University of Hawaii.
- Bickerton, D. 1975. *Dynamics of a creole system*. Cambridge: Cambridge University Press.
- Bickerton, D. 1979. The status of *bin* in the Atlantic creoles. In *Readings in creole studies*, ed. by I. Hancock, I. Polomé, M. Goodman and B. Heine, 309-14. Ghent: E. Story-Scientia.
- Bickerton, D. 1981. *Roots of language*. Ann Arbor: Karoma.
- Binnick, R. I. 1991. *Time and the verb: A guide to tense and aspect*. Oxford: Oxford University Press.
- Boretzky, N. 1983. “On creole verb categories”. *Amsterdam Creole Studies* 5: 1-24.
- Bybee, J. L. and Ö. Dahl. 1989. “The creation of tense and aspect systems in the languages of the world”. *Studies in Language* 13: 51-103.
- Bybee, J. L., R. Perkins and W. Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Byrne, F. and J. Holm. 1990. *Atlantic meets Pacific*. Amsterdam: John Benjamins.
- Chung, S. and Timberlake, A. 1985. “Tense, mood and aspect”. *Language typology and syntactic description*, Vol. III, ed. by Timothy Shopen, 202-58. Cambridge: Cambridge University Press.

- Comrie, B. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Comrie, B. 1985. *Tense*. Cambridge: Cambridge University Press.
- Dahl, Ö. 1985. *Tense and aspect systems*. Oxford: Blackwell.
- Dahl, Ö. 1993. "Review of *Pidgin and creole tense-mood-aspect systems*, ed. by John Singler". *Studies in Language* 17: 251-58.
- Devonish, H. 1978. *Selection and codification of a widely understood and publicly useable language variety in Guyana, to be used as a vehicle of national development*. University of York D.Phil. dissertation.
- Donicic, A. 1954. (with W. Campagne, P. van Daal and others.) *De Creolentaal van Suriname*. Paramaribo: Radhakishun & Co.
- Gibson, K. 1982. *Tense and Aspect in Guyanese Creole: A syntactic, semantic and pragmatic analysis*. University of York D. Phil. Dissertation.
- Gibson, K. 1986. "The ordering of auxiliary notions in Guyanese creole". *Language* 62: 571-86.
- Gibson, K. 1992. "Tense and aspect in Guyanese creole, with reference to Jamaican and Carriacouan". *International Journal of Anthropological Linguistics* 58: 49-95.
- Jaganauth, D. 1987. Predicate structures in Guyanese Creole. University of the West Indies, Jamaica M.A thesis.
- Jaganauth, D. 1988. "Relative time reference in Guyanese Creole: some problems for sentence-level analysis". Paper presented the Society for Caribbean Linguistics, College of the Bahamas.
- Maurer, P. 1993. "Subjunctive mood in Papiamentu". In Byrne and Holm, 243-50.
- Mufwene, S. 1984. "Observations on time reference in Jamaican and Guyanese creoles". *English World-Wide* 4: 199-229.
- Muysken, P. 1981. "Creole tense/mood/aspect systems: the unmarked case?" In *Generative studies on creole languages*, ed. by Pieter Muysken, 181-99. Dordrecht: Foris.
- Nickel, M. and J. Wilner. 1984. *Papers on Sranan Tongo*. Paramaribo: Instituut voor Taalwetenschap (Summer Institute of Linguistics).
- Pollard, V. 1989. "The particle *en* in Jamaican Creole: A discourse-related account". *English World Wide* 10: 55-68.
- Reichenbach, H. 1947. *Elements of symbolic logic*. New York: Macmillan.
- Rickford, J. 1987. *Dimensions of a creole continuum: History, texts and linguistic analysis of Guyanese creole*. Palo Alto, CA: Stanford University Press.
- Sankoff, G. 1990. "The grammaticalization of tense and aspect in Tok Pisin and Sranan". *Language Variation and Change* 2: 295-312.
- Seuren, P. 1981. "Tense and aspect in Sranan". *Linguistics* 19: 1043-76.
- Seuren, P. 1983. "The auxiliary system in Sranan". In *Linguistic categories: Auxiliaries and related puzzles*, Vol. II, ed. by F. Heny and B. Richards, 219-51. Dordrecht: Foris.
- Simons, R.D. 1954. "Het partikel *sa* in het Surinaams". *West Indische Gids* 35: 166-70.
- Singler, J. 1986. "Short note". *Journal of Pidgin and Creole Languages* 1: 141-45.
- Singler, J. (ed.) 1990. *Pidgin and creole tense-mood-aspect systems*. Amsterdam: John Benjamins.
- Spears, A. 1990. "Tense, mood and aspect in the Haitian Creole preverbal marker system". In Singler, 119-42.

- Spears, A. 1993. "Stem and so-called anterior verb forms in Haitian Creole". In Byrne and Holm (eds), 261-75.
- Stolz, T. 1987. "The development of the AUX category in pidgins and creoles: The case of the Resultative-perfective and its relation to anteriority". In *Historical development of auxiliaries*, ed. by M. Harris and P. Ramat, 291-315. Berlin: Mouton de Gruyter.
- Tagliamonte, S. and Poplack, S. 1988. "How Black English past got to the present: Evidence from Samaná". *Language in Society* 17: 513-33.
- Taylor, D. 1963. "The origin of West Indian creole languages: evidence from the grammatical categories". *American Anthropologist* 65: 800-14.
- Voorhoeve, J. 1957. "The verbal system in Sranan". *Lingua* 6: 374-96.
- Voorhoeve, J. 1962. *Sranan syntax*. Amsterdam: North-Holland.
- Wendelaar, W. and Koefoed, G. 1988. "Sa en o in het Sranan". *Oso* 7: 63-75.
- Wilner, J. 1992. The use of *ben* in narrative texts in Sranan Tongo. Paper presented at the conference of the Society for Caribbean Linguistics, Barbados, August 1992.
- Wilner, J. 1994. Non-temporal uses of *ben* in Sranan Tongo. Paper presented at the conference of the Society for Caribbean Linguistics, Guyana, August 1994.
- Winford, D. 1993. *Predication in Caribbean English creoles*. Amsterdam: John Benjamins.
- Winford, D. 1994. "Tense and aspect in Belize creole". Paper presented at the Conference of the Society for Pidgin and Creole Linguistics, Boston.
- Winford, D. 1997. "Property Items and Predication in Sranan". *Journal of Pidgin and Creole Languages* 12: 237-301.
- Winford, D. Forthcoming. "Irrealis in Sranan: Mood and modality in a radical creole". *Journal of Pidgin and Creole Languages*.

Chaos and Creoles

Towards a New Paradigm?

George Lang

1. Introduction

Whether universalist, substratist or some blend, recent work in creole linguistics has operated implicitly within a paradigm which is not only determinist, but linear. The assumption has been that creoles are the effects of causes which can be enumerated and that this causality is driven by rules themselves enumerable, if not for the moment enunciable. In this paper I would like to explore the implications of certain nonlinear models of creole genesis and evolution — without, however, abandoning the principle of determinism, though determinism ought not, chaos theory informs us, be confused with predictability (on this point, see Ruelle 1991, Casti 1995: 85-114).

As Casti, Kellert (1993) and Shermer (1995) make clear, “chaos theory” is something of a misnomer, since chaos is only one of a body of concepts shared by theories of catastrophe, complexification, anti-chaos and simplicity, all of which are connected by their interest in non-linearity. In what follows, I shall nonetheless use “chaos” and “chaos theory” as shorthand for these emerging ideas of order. It is my wager that the concept of non-linearity will shed light on the main paradox creolists have encountered over the years: the strange continuities within the discontinuity of creoles. These are usually described in terms of widespread commonalities among tense-mood-aspect marking systems, patterns within noun phrases, the use and forging of reflexives, and serial verb behavior, but are also related to the consistent presence of dialect continua in contemporary creole societies. Alain Kihm reminds us in the course of his ebullient embrace of eclecticism in the introduction to *Kriyol Syntax* (1994: 2) that “the scientific method is nothing

more (nor less) than educated curiosity and debate” (1994: 10). The present preliminary inquiry into what might be a new paradigm is offered as nothing more or less than that, but is in addition inspired by Kellert’s recommendation that we “initially construe ‘chaos theory gives us understanding’ to mean nothing much more than that it helps us to think about and respond to some aspects of the world in interesting or useful ways” (1993: 80).

At the outset I should state that mathematicians are leery of applying precepts of chaos theory to the social realm, warning that such extrapolation is potentially metaphysical. In particular, Kellert warns that “research [into chaos theory] is still technical and mathematical in nature” and “often yields precise numerical results” (1993: 4). He later denounces the “fashionable rhetorical co-optation” of the concept of chaos, especially in the study of culture (1993: 77), *cautio* which should be kept in mind.

Increasingly, doubt has been raised that creoles actually constitute a class of language which can be defined on purely typological grounds, from which it would follow that we must know the social history of a language in order to baptize it a creole. Creoles also may not belong to a sharply defined set but to a Wittgensteinian “family” composed in terms of overlapping criteria of resemblance, as Mufwene has suggested (1993: 133), in which case their pedigree would remain a matter more of speculation about common facial features than of DNA testing, as it were. Also bearing on this point is Kihm’s admission that it is basically irrelevant to his study that Guinea-Bissau Kriyol is a creole, since it is “presently...an ordinary representative of the natural class ‘human language’” (1994: 2). All of these perspectives converge towards skepticism first about the discrete and autonomous existence of creoles (or the more inclusive category, pidgin-creoles or PCs), and second about the prospects of defining them satisfactorily.

If creoles or PCs are to be defined non-tautologically, that is as something “besides the object of PC studies” (Byrne & Holm 1993:2), they must be qualified as belonging to a restricted set of systems. That set could be defined in terms either of common genetic dynamics (universality, substratist, etc.), or of other post-genetic features of those language systems. I can thus see two ways in which chaos theory might fruitfully apply. First, there are good intuitive reasons to compare creole genesis with the onset of “turbulence”, which is the harbinger of chaotic order. The implications of this comparison are discussed in sections 2 and 3 below. A second line of thought is advanced in section 4: leaving aside the issue of creole genesis, chaos theory might offer

a frame of reference with which to identify contemporary creoles *qua* creoles. For example, if “linguistic turbulence” can be shown to be a continuing feature of contemporary creoles, which is certainly implied when creoles are framed in a continuum and even more when they are said to decreolize, then there might be a typical profile of “fractality” with which to typify them. In both cases, my remarks should be considered merely exploratory.

2. Some Definitions

Chaos, though understood to represent deeper forms of order, is associated with instability and aperiodicity. Such irregularity is an attribute of mathematical models of chaos but is also attested in the physical world, first and foremost during the onset of hydraulic turbulence but, on a larger scale, in any complex system, like weather, where the outcome of events is sensitive to initial conditions. The classic image of sensitivity to initial conditions is in fact meteorological, and takes the form of asserting that the weather at this moment in, say, San Diego could not have been predicted a few weeks ago, simply because the unrecorded effects of a butterfly’s wings in Guyana would have been multiplied incalculably, even had we some way to gauge eddies swirling off those lepidopterian wings, or had thought to. It remains true, though, that once a Pacific storm is blowing in, those tiny wings would flap in vain.

Kellert defines chaos theory as the “*qualitative study of unstable aperiodic behavior in deterministic nonlinear dynamical systems*” (1993: 2; my emphasis). It is worth lingering on and expanding these terms of reference.

Dynamical systems, for Kellert, combine formal descriptions of “the instantaneous state of a physical system” with rules for “transforming the current state of description into a description for some future, or perhaps past, time”. In the simplest dynamic system, a single variable is altered and the *ceteris paribus* principle that all other things remain equal produces apparent proof of cause and effect. The problem in any system sensitive to initial conditions is that all other things are never equal unless *all* the initial conditions are accounted for. This predicament is endemic to the human sciences, where systems cannot be easily isolated, where initial conditions reside in the distant past and where (positivistic) rules for “transforming” the modeled “instantaneous state” into another past or future one are grossly inadequate.

Nonlinear, in Kellert’s terminology, denotes having “algebraic or other

more complicated functions of the system variables..., expressions such as x^2 or $\sin(x)$ or $5xy$ " (Kellert 1993:3). The difference between nonlinear and linear equations is that latter are "closed" and allow one, two or three variables, examples of which the linear equation in one variable $ax + b = 0$, or that in two variables x and y : $ax + by + c = 0$ (Clapham 1990: 105). This mathematical distinction between linear and nonlinear is not terribly germane to my discussion, since creolists are far from able to produce formulas of either variety to account for language contact and change. Yet whatever does go on during creole genesis and during the subsequent evolution of a creole, the variables at play are certainly more than three and are rather complex, to say the least. In any event, the appeal of chaos theory to historians and social scientists follows from the recognition that "accurate dynamical models of history, were we able to produce them, would be nonlinear" (Reisch 1995: 58); or again: "simple linearities would be surprising in anything as complex as a person or a social system" (Dyke, in Shermer 1995: 64). Similarly, were some equation to apply to the dynamic system of a creole, it would emphatically be nonlinear, containing a plethora of variables.

As a kind of thought experiment, let us imagine what the multiple variables of a nonlinear model of creole evolution might include. The external ones are already numerous: for example, the ratio of various populations during creole genesis (a classic exposition of which is Baker & Corne 1986); the degrees of hegemony and subservience and heterogeneity or homogeneity of each population (different at each site of genesis); the number and nature of the substrate languages involved and/or the dialectal complex of the superstrate; the extent of sociological expansion of the nascent creole at or shortly after the moment of genesis. At some mediate point between external and internal factors is the particular typological dynamic of the creole in question, usually thought to be either exogenous or endogenous (per Chaudenson [1979:21], those which arise when the initial population of initial speakers has been displaced into a new environment versus those which develop *in situ* around an implanted foreign population). Internal variables are even more forbidding. Typical creole features have been attested on all levels of contemporary linguistic analysis: phonological, morphological, syntactic, semantic. Further difficulty results from the recent consensus that creoles are only one sub-set of contact languages (Arends, Muysken & Smith 1995; Thomason 1997), and that contact produces language types distinct from creoles *per se*, for example the Canadian Prairie French-Cree mixed language

known as Michif, which Bakker & Papen claim is “unique...in the world” (1997: 356). The specter of these spectra is awesome for anyone seeking to situate a given creole within the array of deterministic factors which have shaped it.

Qualitative is an ambiguous term since it is often opposed to “quantitative.” Chaos theory is oriented towards precise mathematical results and is thus in the conventional sense of the word quantitative, “number-crunching”. For Kellert, however, there is a difference between quantitative closed-form solutions, such as predicting when three planets traveling in elliptical orbits will align, or a comet will impact Jupiter, and *qualitative* study, which “would be more interested in discovering what circumstances will lead to elliptical orbits as opposed to, say, circular or parabolic ones”. Dynamical systems theory asks questions like: “what characteristics...*all* solutions of [a] system ultimately exhibit” (Kellert 1993: 4).

This statement decidedly reflects a step back to a higher degree of abstraction than any closed-form solution might entail. Kellert is implying that (mathematical) solutions themselves be submitted to mathematical manipulations, and such *mise-en-abyme* is precisely what has enabled chaos theoreticians to produce all those pretty graphics, which have been generated by recursive iteration of formulas. To be sure, creolists are far from having worked out the “solutions” to the numerous languages they have made their own. But if there is a valid autonomous agenda for creole linguistics (i.e. if creolists are not just linguists with a relish for an open-ended assortment of exotic socially marginal languages), then that agenda must be to discover which circumstances produce creoles, rather than, say, pidgins or mixed languages or any among the natural class of human languages. It is crucial to distinguish, however, between the relatively simple operations of defining terms and finding examples (here is a pidgin, there is a creole), and the sort of recursive abstractions which are proper to chaos theory.

For Kellert, chaos theory is the set of systems theory which focuses on the qualities of physical “behavior which is unstable and aperiodic..., systems which never settle into a form of behavior that resists small disturbances. A system marked by stability, on the other hand, will shrug off a small jostle and continue about its business” (Kellert 1993: 4). The distinction in physical sciences between stable and unstable systems and their differing degrees of predictability is often easy to make. To borrow Shermer’s example adumbrated above, science could predict with impressive accuracy the moment the

comet Shoemaker-Levy 9 would crash into Jupiter, but the results of that collision were wholly unknown in advance, not only because of lack of knowledge about the structure of Jupiter, but by virtue of the red planet's nature as a meteorological complex — i.e. an unstable system (1995: 69). The challenge is to transpose this distinction into that between stable and unstable language systems.

Sensitivity to initial conditions refers not merely to the fact that the past and present of complex systems cannot be adequately measured, but to a second important facet of chaos. Obviously, the earlier an event occurs in any deterministic sequence, the greater its effect; hence the need to have the most complete possible account of early or initial conditions. It happens, though, that systems collapse. Usually, these moments of catastrophic change are triggered by minute factors, proverbial straws which break the camels' backs. Furthermore, a phase of utter unpredictability emerges from catastrophe. It is as if a multitude of new parameters of the subsequent order of things are being reset. There arises a welter of new "initial conditions" to which later events must be sensitive. In chaos theory, such post-catastrophic moments are called *bifurcation points*. In human history, war provides a compelling site of bifurcation points and hence chaotic determinism. By way of illustration, Shermer cites historians' commonplace adage that for want of a horseshoe nail, the kingdom was lost, but adds that once the kingdom has collapsed, ordinary necessity rules again, and "100,000 horseshoe nails will not help a bit" (1995: 70).

Although it is possible that creolization is more gradual than abrupt (cf. Arends 1993), it is an implicit precept of most creolists that once a creole has crystallized or jelled, it becomes more or less invulnerable to deep restructuring (though it may decreolize). In other words, the above list of external and internal dynamic variables at play during creole genesis may be thought to apply largely during the transformative catastrophic phase of a creole's history. Afterwards, these variables are not causal, rather mundanely descriptive of a steady state.

3. "Linguistic Turbulence" and Entropy

Readers will have understood that I am proposing that creole genesis may be one site of chaotic determination in history. Even endogenous creoles can be thought of as arising during conditions of linguistic turbulence, but so-called

“abrupt” or “radical” creoles, per Thomason and Kaufman (1988: 166) are most readily pictured as transpiring when the steady flow of language transmission from one generation to the next is interrupted or fractured, and hence as consisting of innumerable bifurcation points at which normal linguistic necessity breaks down. Ideally, we would be able to devise some hands-on experiment with creoles to explore those bifurcation points, much like Edward Lorenz did with his drips dropping into a waterwheel, or the biologist Robert May did for bust-and-boom wildlife population cycles (Gleick 1987: 30,71). Unfortunately, proofs consonant with chaos theory require reams of data, which we all admit to lack about creole genesis. To mix metaphors, but by way of allusion to a possibility haunting research on early creole history, genesis may be a sort of black hole from which information, at least the sort to which we are accustomed, cannot escape.

From the point of view of those seeking continuities across the great divide of genesis, the worst implication of this state of affairs is that if genesis is chaotic, then the path of any lexical, syntactical or phonological trait would remain indeterminate, in the same way that the postlapsarian position of two molecules of water in proximity at the top of a waterfall cannot be predicted; or, more to the point, vice versa: their position above cannot be known from below.

I am not arguing in favor of complete agnosticism about early creole history, only that the frame or scale in which knowledge is possible may be much broader than we would like. Although the relative position of the two aforementioned molecules in a waterfall is beyond even the hardest hydraulic science, we would still know without a shadow of a doubt which waterfall we are dealing with. In most cases, we do know which languages we are talking about (though perhaps not always). Moreover, every creole has a post-genetic history which, like that of any natural language, is subject to change and shift. If creoles are indeed chaos-born, there would nonetheless be some telling traces in post-chaotic history, these deriving from their identity as a special class of languages abruptly created in crisis, although, once jelled, evolving like any other language in contact with others.

This may well help explain why so many “non-creole” features are found both in comparatively ancient creoles, like Capeverdean Crioulo, or ones with significant interal diversification due to subsequent language contact, such as, again, Capeverdean, or Papiamentu. By the same token, unless creoles are a specific product of recent human history since, say, 1500, we must presume that some must have been produced prior to that date and evolved normally

(“normally” including contact with other languages) long enough to be indistinguishable from other languages, at least on the basis of the information they bear in their structures. The alternative is also intriguing: that the classic Atlantic creoles were an exceptional one-time creation of human history.

Another potential application of the chaos scenario to creole development is observed by McWhorter (1998), who among others claims that creoles begin by exhibiting a degree of structural regularity unusual among spoken languages, but quickly begin evolving entropically, their derivations becoming “opaque”, less “semantically regular”: in other words, they become clotted and encrusted with the forms of normal language. *Entropy* has a precise meaning in thermodynamic theory, and has also acquired a somewhat metaphoric sense in communication or information theory, as the irreversible tendency of any isolated system “to slide toward a state of increasing disorder” (Gleick 1987: 257). Disorder is synonymous with randomness and chaos, the latter term here in the colloquial sense of unpredictability. Transposed into creolistic phraseology, then, “entropy” might be said to refer to the above-mentioned increasing lack of transparency in derivations obtained as usage smudges and pastes over the semantic regularity produced during genesis.

In ordinary (non-creole) conditions of language change, the measurement of entropy in this sense is meaningless, since the history of any language stretches back into the fog of time. To be sure, grasp of deep regularity is one goal of contemporary linguistics, but if deep forms do exist in language, they ought to be equally deeply hidden in all natural languages. If deep forms in creoles lie “closer to the surface”, overlain by a relatively shallow layer of entropic accumulation, then creoles would offer privileged windows into the nature of those deep forms. This was, after all, the gist of Bickerton’s Language Bioprogram Hypothesis (LBH).

Yet one need not espouse the LBH to conceive of creoles as having originary moments in history and hence potentially bearing evidence of entropy. One reason why such entropic opacity faces insurmountable problems of measurement is that creoles are not isolated systems. They are in fact prototypical non-isolated systems, always entangled in language contact above and beyond the norm. This common characteristic of creoles (which I have called “entwistedness” in a study of literature in creoles [Lang 1999]) leads to the second general approach we might take in applying chaos theory to creoles: seeking a typical signature of chaos in their synchronic state as languages in intense contact.

4. A Fractal Quotient for Creoles?

Before exploring this avenue, I need to open a parenthesis on the numbering of dimensions in chaos theory.

Chaos theory initially arose out of “number experiments” (Kellert 1993:92), which is why hard physicists resisted its applications for so long. The seahorse tail swirls of the Mandelbrot set, that quintessential image of fractal self-similarity, were, for example, obtained by plotting the results of iterated formula of complex numbers against coordinates on a computer screen grid, constructing what is called phase space in which all possible states of a system are represented in points arrayed by any number of dimensions (usually, for practical reasons, in two or three, with segmental dissections in one). Gleick’s description of how an iteration loop is plotted in phase space is very accessible (1987: 231-2), and conveys the wonderment felt during the 1980s as ever increasing computing power graphically displayed hidden dimensions of order within numbers themselves, in particular the similarity of pattern across scale which is the stamp of fractality. (Pickover 1990 supplies a myriad of similar operations for those inclined and equipped to pursue recursive experiments productive of fractality.)

Self-similarity and fractality are contingent upon perspective, upon scale. To take another example from Gleick (1987:97), a ball of twine seen from a great distance constitutes a mere point, with zero dimension; closer up, it is a woolly three dimensional ball; from closer still the twine is revealed to be unidimensional (that is, a point in it is situated in terms of its longitudinal pole, though wrapped around itself within three dimensions), with various two-dimensional surfaces dissecting the entwined tubes defining multitudinous planes of bidimensionality, etc. Depending on the frame of reference and the unit of measure, then, dimensions can therefore be considered fractional in nature: one can be at various points “in between” the three spatial dimensions, which are conventionally enumerated by integers. Similarly, the length of a shoreline depends on the scale with which it is measured: the steps of a hiker’s stride, the clambering of an ant, microns, or angstroms. Furthermore, the finer the instruments of measure, the more detailed the data image obtained. This is apparent to anyone involved in empirical research, but upon detailed parsing, some data images are “self-similar,” their features remaining constant through successive magnifications, like coastlines or mountains. Self-similarity across changes of scale is a hallmark of fractality, said to exist when a profile

of irregularity is constant over different scales. It is in fact possible to attribute a given fractal dimension (D) to objects like the shoreline of Britain ($D = 1.26$), or to phenomena like speech waveforms, which according to Pickover again, are statistically self-similar at sentence time scales, and have a fractal dimension of 1.66 (1990:58-61).

The empirical validation of chaos consists of two steps: first the elaboration of a recursive formula corresponding to the mathematical patterns which show fractality, and second the discovery of actual phenomena which fit those formulas. Gleick cites examples from fluid dynamics and meteorology, but also biomedicine and economics. For reasons explained above, there is and will remain insufficient data to subject creole genesis itself to such operations. On the other hand, analysis of *contemporary* creole speech situations might alone yield a sufficient bulk of data. The specificity of creoles would then be shown to be their having typically fractal patterns. The implicit premise here is that creoles do constitute a class in and of themselves, with a (fractal) logic of their own. Is it possible to plot variables within creole continua not against properties and items, but in phase space?

Implicational scaling of variation within a creole continuum is one of the most powerful descriptive tools we possess, and one of the rare quantitative ones. For illustrative purposes, let us consider Rickford's matrix representing a "continuum-like transition between creole and standard pole" (1987: 19 Table One).

As Rickford himself observes, such matrices are rough and arbitrary: data always seem to allow for exceptions (1987: 17). Even grammatical categories shade off from one to the other much like, incidentally, the color red did during the thought experiments of Feigenbaum, another pioneer of chaos theory (Gleick 1987: 164). Nonetheless, as long as the primary terms of implicational scaling are based on integers and simple fractions, a "lower" non-chaotic degree of order is likely to appear. Take, for example, the first feature Rickford attributes to Guyanese Creole: variations between palatized and non-palatized velar consonants, *gy* in "car" and "girl". This is solid phonology, but the phonetic basis of this distinction is rather blurry (and, as mentioned above, wave forms themselves appear to be fractal). Suppose, instead, that the actual phonetic variation falls in between the integral dichotomies (X versus non- X) habitually used to measure them for phonemic purposes. If more numerous points were plotted and then subjected to more powerful and less linear models, we would, perhaps, find ourselves in a

Table 1. Linguistic features of four texts representing a continuum-like transition between creole and standard poles (adapted from Rickford 1987: 19)

Text	gy	d	a V	V+Ø	aa	dem	na V	wan
20	1	1	1	1	1	1	1	1
21b	1	1	1	1	1	1	1,2	2
15c	2	1,2	2,3	1,2	1,2	2	2,3	2
30	1,2	1,2	2,3	2	2	—	3	2

Index:

- gy: 1 = a palatalized velar consonant in words like “car” and “girl”
 2 = a nonpalatalized velar
- d: 1 = initial stop in words like “this”
 2 = initial interdental fricative
- a V: 1 = continuative aspect marked by preverbal *a*
 2 = continuative aspect marked by verb + “-ing”
 3 = continuative aspect marked by “be” plus verb + “-ing”
- V+Ø: 1 = bare verb-stem used to mark past tense with nonstative verbs
 2 = “-ed” suffix or stem change used to mark past tense with nonstative verbs
- aa: 1 = a low unrounded aa in words like “all”
 2 = a rounded *oh*
- dem: 1 = *dem* is third-person plural subject pronoun
 2 = *dee* is third-person plural subject pronoun
- na V: 1 = preverbal *na* or *no* as negator
 2 = preverbal *en* as negator
 3 = inflected form of “do” plus “not” as negator
- wan: 1 = *wan* as indefinite article
 2 = “a” as indefinite article

situation similar to that encountered when fluid dynamics became embroiled in the number experiments of early theorists of chaos. To be sure, the implications of this kind of procedure are not specific to creolists, but to linguists as a whole, ultimately calling into question the distinction between phonemics and phonology.

At first sight, grammatical categories, like those between continuative aspect marked by preverbal markers (*a*) and that marked by verb+ *-ing*, seem much more naturally distinct (the third and fourth features in Rickford’s matrix). Rickford’s own data shows, however, that vacillation between them is present in some dialects or, more accurately, texts. Such vacillation is doubtlessly contextually determined, and perhaps could be situated in their contexts and further quantified. This degree of quantification would require many more variables than are possible in a polar or bidimensional matrix. The familiar contrast between digital and analog paradigms would be one way to classify this complexity, but chaos theory itself remains digital, in the sense

that it partitions analogical continua into discrete though virtually infinite points of measurement. In other words, were measurement of this or any other creole continuum more precise, or at least more variegated, these continua could perhaps be made to yield the kinds of variability found in fluid dynamics. Again, this methodological issue is not specific to creolistics, but concerns general linguistics.

In his discussion of the Guyanese Creole continuum, Rickford concludes that it is “*theoretically* desirable to attempt to restate multidimensional analyses in unidimensional terms” (1987: 38; my emphasis). I have underscored the adverb in Rickford’s statement because I take it to be pivotal. When I first encountered it, I remembering wanting to argue that the word “practically” ought to replace “theoretically,” since it is pragmatically difficult first to gather, then to configure and represent data like this without recourse to simplification, whereas it is theoretically possible to imagine not only vastly more dimensional matrices, but an infinitely dense scattering of idiolectal points in any given hypothetical continuum. Taking a hint from the organizers of the 1996 Society for Pidgin and Creole Linguistics sessions in San Diego, I now see the practicality of Rickford’s word “theoretically.” For the moment it is better to say merely that it is *metatheoretically* possible to question not merely unidimensionality or multidimensionality, but integral dimensionality *per se*. The integrality of index features of any matrix, perhaps necessary for reasons of concision, is arbitrary; a very large number, indeed an infinity of other fractional features must be presumed to lie in between or among them, features which might be fractal. The result might be called *hyperdimensionality*.

One “soft” (non-numerical) reading of such hyperdimensionality would follow the precepts of Bakhtinian translanguistics by applying dialogical heteroglossia to continua of creole speech acts (see Garrett 1994). *Heteroglossia* is, according to Bakhtin (1981:428):

The base condition governing the operation of meaning in any utterance. It is that which insures the primacy of context over text. At any given time, in any given place, there will be a set of conditions — social, historical, meteorological, physiological — that will insure that a word utterance in that place and at that time will have a meaning different than it would have under any other conditions. All utterances are heteroglot in that they are functions of a matrix of forces practically impossible to recoup, and therefore impossible to resolve. Heteroglossia is as close a conceptualization as is possible of that locus where centripetal and centrifugal forces collide; as such, it is that which a systematic linguistics must always suppress.

Although I believe the Bakhtinian model to be heuristically useful, I am stymied by the problem that any qualification of a dialogical act is itself dialogical, and that the precise constellation, or indeed nebulae of utterances speakers organize into dialects and idiolects is conjectural, accessible to the artistic or critical imagination, but too squishy for the taste of linguists. (Such a mushy compound of critical and chaos theories as applied to Curaçao and Papiamentu can be found in Lang 1997.) In any event, if verified, the fractality of creole continua would be very much in accord with conceptions of creoles as clusters of speech acts, a line of thinking going back through Garrett, Carrington (1992), and Le Page & Tabouret-Keller (1985). A radical statement of this position was uttered by Le Page as long ago as 1967: “There is,” he said, “no such thing as a language except insofar as the idiolects of two or more people overlap” (in Taylor 1977: 225). The nonlinear dynamics which I see emerging from a tentative application of chaos theory to creolistics is not quite so radically nominalist, since the particulate distribution of idiolects in any creole continuum may resolve into patterns, figures and designs which display considerable dimensions of order, though dimensions more complex than those projected in matrices of creole continua that are predicated upon integers, or whole numbers. Note that it was precisely when he grasped that dimensions were fractional, i.e. inter-integral, that Benoit Mandelbrot conceived of his fractal geometry, which he coined from Latin *fractus*, broken (Gleick 1987: 98): here it is hard not to think of Schuchardt’s *Radebrecher*, the wheel of torture upon which syntax was broken down into fragments.

5. Conclusion

I had hoped that one trade-off of detecting chaos within creoles would be the following: given the regularity of their irregularities, creoles would have characteristic topologies measurable in fractal terms. At this stage, the best I can offer is the suggestion that topologies yet to be sought might better project finer images of the higher order within the chaos of creoles. I am deliberately using visual, graphic and pictorial wording. As Kellert observes, one particularity of chaos theory is its use of new geometries to describe, but not necessarily to *explain* order (1993: 106). An analogous procedure in creolistics would be not to account for the correspondences between tense-mood-aspect systems in creoles as the result of some deep cause(s), but rather to map it in all of its

intricacy and then *display* the patterns thereby captured; in Kellert's words: not law(s), but order (1993: 111).

Creolistics is the study of hybrid, heterogeneous and marginal language contact situations. Hence, the potential impact of chaos theory, which was devised in order to account for experimental findings which had been systematically ignored in mainstream physics. Kellert even claims that chaos theory emerged in contradistinction to a prevailing "clockwork hegemony" in physics, with its predisposition towards "microreductionist, deductivist, and synchronic" methods (1993: 86). I shall not follow Kellert into the assertions of his last chapter, buttressed by allusion to feminist historians of science, that repression of "turbulent" and "chaotic" phenomena reflects a gendered agenda in science. It might well be, though, that whatever position one occupies along its doctrinal and methodological spectrum (or shouldn't I say: within its multidimensional ideational space?), creolistics is inherently a theory of chaos in linguistic science.

References

- Arends, J. 1993. "Towards a Gradualist Model of Creolization". In Byrne and Holm 1993, 371-380.
- Arends, J., Muysken, P. and Smith, N. 1995. *Pidgins and Creoles: an Introduction*. Amsterdam: John Benjamins.
- Baker, Philip and Chris Corne. 1986. "Universals, Substrata and the Indian Ocean Creoles". In *Substrata versus Universals in Creole Genesis*, P. Muysken and N. Smith (eds), 163-83. Amsterdam: John Benjamins.
- Bakhtin, M. M. 1981. *The Dialogic Imagination*. (Ed. by Michael Holquist, trans. by Caryl Emerson and Michael Holquist.) Austin: University of Texas Press.
- Bakker, P., and Papen, R.A. 1997. "Michif: A Mixed Language based on Cree and French". In Thomason 1997, 295-363.
- Byrne, F. and Holm, J. (eds) 1993. *Atlantic Meets Pacific*. Amsterdam: John Benjamins.
- Calpham, C. 1990. *Concise Oxford Dictionary of Mathematics*. Oxford University Press.
- Carrington, L.D. 1992. "Images of Creole Space". *Journal of Pidgin and Creole Languages* 7.1:93-99.
- Casti, J.L. 1995. *Complexification*. New York: HarperCollins.
- Chaudenson, R. 1979. *Les Créoles Français*. Évreux: F. Nathan.
- Garrett, P.B. 1994. "Conceptualizing Linguistic Variation in Creole Environments". Paper presented to the joint meeting of the Society for Pidgin and Creole Linguistics and the Society for Caribbean Linguistics, University of Guyana.
- Gleick, J. 1987. *Chaos: Making a New Science*. New York: Penguin.
- Kellert, S.H. 1993. *In the Wake of Chaos*. Chicago: University of Chicago Press.

- Kihm, A. 1994. *Kriyol Syntax: The Portuguese-based Creole Language of Guinea-Bissau*. Amsterdam: John Benjamins.
- Lang, G. 1997. "Papiamentu, the Caribbean Paradigm and Cultural Fractality". In *Latin American Postmodernisms*, R.A. Young (ed.), 87-102. Amsterdam: Rodopi.
- Lang, G. 1999. *Entwisted Tongues: Comparative Creole Literatures*. Amsterdam: Rodopi.
- Le Page, R. and Tabouret-Keller, A. 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.
- McWhorter, J.H. 1998. "Identifying the Creole Prototype: Vindicating a Typological Class". *Language* 74: 788-818.
- Mufwene, S. 1993. "Are There Possessive Pronouns in Atlantic Creoles?" In Byrne and Holm, 133-143.
- Pickover, C.A. 1990. *Computers, Pattern, Chaos and Beauty: Graphics from an Unseen World*. New York: St. Martin's Press.
- Reisch, G. 1995. "Scientism without Tears: A Reply to Roth and Ryckman". *History and Theory* 34.1:45-58.
- Rickford, J. 1987. *Dimensions of a Creole Continuum*. Stanford: Stanford University Press.
- Ruelle, D. 1991. *Hasard et chaos*. Paris: Éditions Odile Jacob.
- Shermer, M. 1995. "Exorcising Laplace's Demon: Chaos and Antichaos, History and Metahistory". *History and Theory* 34.1:59-83.
- Taylor, D. 1977. *Languages of the West Indies*. Baltimore: Johns Hopkins University Press.
- Thomason, S.G. 1997. *Contact Languages: A Wider Perspective*. Amsterdam: John Benjamins.
- Thomason, S.G. and T. Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.

Wh-words and Question Formation in Pidgin/Creole Languages*

J. Clancy Clements and Ahmar Mahboob

1. Introduction

1.1. *Goal of the study*

In this study, wh-words are employed as a diagnostic of the relationship between several Portuguese-based and French-based creoles on the one hand, and various Spanish- and English-based creoles on the other. The second goal of the paper will be to demonstrate a principal argument that wh-words demonstrate the role of *mutual linguistic accomodation*, as defined by Thomason & Kaufman (1988), in creole genesis.

1.2. *Wh-words in pidgins and creoles*

In classifying creoles and reconstructing their origins on the basis of synchronic aspects of a creole's grammar, we must ascertain that the aspects in question trace back to the time of genesis, rather than being later developments. In this light, Muysken & Smith (1990) point out the following interesting features exhibited by question words:

- a. Question words have a central status in the lexicon as function words (p. 883)
- b. Question words are normally less susceptible to replacement amidst processes of historical change than content words (ibid. 883)
- c. At least in 18th century Surinam creoles, question words do not support the gradualist view of creolization (ibid. 900).¹

On the basis of these observations, we will assume in this study that the

question word system generally develops early in the formative stages of a pidgin or creole, and subsequently tends to be less subject to change than content words. (We will see, however, that there are exceptions, due to factors such as decreolization [Daman Creole Portuguese], relexification [Papiamentu, Palenquero], etc.)

1.3. *Wh-word constructions cross-linguistically*

The placement of *wh*-words in questions involves what are commonly called *partial questions*, in which the speaker and hearer share knowledge regarding a given proposition. This knowledge is presupposed, but the speaker lacks one element of the proposition, the one referred to by the interrogative form. The missing element, which can be the subject, object, verb, predicate, adverb, time, place, manner, etc., is usually considered the focus of the utterance (Givón 1990:793). Indeed, *wh*-word positions across the world’s languages reflect their focused nature. As in the case of focused elements, the most common position of a *wh*-word is clause-initial. Even in verb-initial languages such as Tagalog, Malagasy, or Jacaltec, one of the few elements that can be placed clause-initially is the *wh*-word. For SOV languages, the unmarked position is not clause-initial, but rather in preverbal position, that is, *in situ* (Givón 1990:800), as summarized in Table 1:

Table 1: *Unmarked position of wh-words according to language type*

	more frequently	less frequently
SVO Languages:	clause-initial	<i>in situ</i>
VSO/VOS Languages:	clause-initial	<i>in situ</i>
SOV Languages:	<i>in situ</i> (i.e. preverbal)	clause-initial

1.4. *Wh-words and universals: bimorphemic question words*

As to the reason for the bimorphemic nature of *wh*-words in pidgins and creoles, Muysken & Smith (1990) propose that bimorphemic *wh*-words are semantically more transparent than monomorphemic ones. For example, if we compare mono- and bi-morphemic reflexes of the same concept, we see that the latter is much more systematic and would conceivably be easier to learn, entailing merely the combination of *what* with a noun already known to the

speaker and hearer as core vocabulary:

(1)	who	what person
	what	what thing
	when	what time/day/hour
	where	what place
	why	what purpose
	how	what manner
	how much/many	what quantity
	which	what person

Bimorphemic wh-words also fit Seuren & Wekker's (1986) criteria for semantic transparency:

- a. Uniformity: Maximal uniformity of treatment of semantic categories
- b. Universality: Minimal reliance on rules or rule types that are highly language-particular
- c. Simplicity: Minimal processing

All the bimorphemic wh-words in (1) are *uniform* in combining one question word *what* with one of a number of frequently used nouns; the system does not rely on language-specific rules and is thus *universal*, and finally is simpler than the monomorphemic paradigm in requiring the acquisition of fewer lexical items.

1.5. *Wh-words and substrate influence: typological distance and mutual linguistic accommodation*

In Portuguese-, French-, Spanish-, and English-based creoles, we find that the structure of questions often reflects shared structures among the languages in contact, what Thomason & Kaufman (1988) call *mutual linguistic accommodation*.

The formation process of a pidgin or creole necessarily involves the creation of a new linguistic system by speakers of two or more linguistic communities. To illustrate, let us take as examples Mauritian French Creole and the Portuguese-based Sãotomense, Angolar, and Principense, spoken on the west coast of Africa.

At the time of the Portuguese and French colonization of West Africa and the Indian Ocean islands, the varieties of Portuguese or French which captured

Africans were exposed to can be assumed not to have been standard; at the very least, they were simplified and, in the case of Portuguese, possibly pidginized. Needing a language for communication with others speaking various native languages, captives shifting to these simplified or pidginized varieties made guesses about what their interlocutors would understand as they tried to talk to them. The grammar of the emerging pidgin or creole would be a reflection of the shared “right” guesses (Thomason & Kaufman 1988:152-3). Since there was a target language (henceforth TL) in the cases we are referring to, slaves would learn the vocabulary of the TL as it was available; at the same time, they would transfer structures from their own languages into the emerging contact language.

In a contact situation such as on São Tomé or Mauritius, if the native languages involved and the TL had a structure in common, it is likely that the creole that emerged from such a contact situation would also contain that structure. *Mutual linguistic accommodation* entailed that during the acquisition process, an African or Malagasy speaker trying to learn the variety of Portuguese or French to which he had access would adapt the structures of his language to those of the TL so long as there was a fit. The factors affecting mutual linguistic accommodation in pidginization and creolization are (Thomason & Kaufman 1988:157):

- a. Access to the TL
- b. Motivation on the part of the shifting speakers to learn the TL in a context of creolization.
- c. The typological distance that separates the languages in a given contact situation
- d. The tendency to use universally unmarked features in the creolization process.

The order of these factors is important, given that there is no question of linguistic interference if there is no access to the TL, or if the attitude of the speakers precludes it (cf. T&K:35). Where attitudes support contact and there is access to the TL, typological distance becomes a factor. If there is little typological distance between the languages in a contact situation, the likelihood of the contact languages incorporating structures from those languages is increased. In contrast, if the languages are typologically very distinct, shifting speakers will tend more to fall back on universally unmarked features. Thus there is an inverse relation, expressed by Clements (1992), between typological distance and mutual linguistic accommodation:

Where there is a good fit between structures of the languages of the contact situation, the probability that the emerging creole will contain these structures is increased. Where typological fit is less, i.e. where there is greater typological distance between the languages of the contact situation, the probability of finding structures of the contact languages in the emerging creole is decreased.

2. Wh-words and Question Formation in Portuguese-based Creoles

2.1. The Portuguese-based creoles of West Africa

2.1.1. Wh-word inventories

The wh-words from these five languages, given in Table 2, reveal that although the lexical material is from Portuguese, exact correspondences are often lacking.

Table 2. Comparison of wh-words in West African Portuguese-based creoles

	<i>quem</i> 'who'	<i>que</i> 'what'	<i>quando</i> 'when'	<i>onde</i> 'where'	<i>porquê</i> 'why'	<i>como</i> 'how'	<i>quanto</i> 'how many'	<i>qual</i> 'which'
KV	ken (ki) /kena	kus'e	kānd	undi	pamo- di k'	modi ki	kantu, kònte	kal, ki
Kr	kin ku (bakin)	ke ku	kalhora ku/ kwandu ku	nunde ku	pabya dike ku, keku	kuma ku	kantu... ku	kal
Pr	nĩgé ki...â	kwá...â	(kwa) zo ki...â	kumí... â	kwe māda ki	mɔdi / mɔ ...â	kwātu ki...â	nĩgé/ kwá...â
Ang	ngê ngêi	kwa, kwai (ma/ki)	dia kutxi, ola kutxi	a, andji (ma)	ra kwai, kwa ma...ra	m'ma (ma)	kantu	kutxi
ST	keŋé ku	kwa ku	(ku) ola (ku)	'nāji	kwa mādá	mɔd(i) (ku)	k(w)ātu ku	
Ann	kenge	xa	ke ola	xama	xafe	xamá	xántu	kixi

(KV = Kabuverdianu [Cape Verdean], Kr = [Guinea-Bissau] Kriyol, ST = Sãotomense, Pr = Principense, Ang = Angolar, Ann = Annobonese; data from Günther 1973, Ferraz & Valkhoff 1975, Kihm 1994, Bartens 1995, Maurer 1995)

Although the data do not present the complete picture, they reveal that the creoles that correspond most closely to Portuguese are Kriyol and Kabuverdianu. Kriyol has 78% (6/8) of Portuguese’s wh-words, while Kabuverdianu has 63% (5/8). In contrast, Sãotomense displays only 43% (3/7), Angolar 25% (2/8), and both Principense and Annobonese a mere 13% (1/8) rate of correspondence.

This suggests that the Upper Guinea creoles formed, at least in part, independently of the Gulf of Guinea ones, which is further demonstrated by a quantitative comparison of the wh-words the former share with the latter versus those that the latter share among themselves.² For example, Table 3 suggests that the Upper Guinea creoles formed, at least in part, independently of the others: they share 63% (5/8) of their wh-words, but much lesser percentages with the Gulf of Guinea creoles.³

Table 3. Percentage of wh-words shared between Upper Guinea and Gulf of Guinea Portuguese creoles

	Kabuverdianu	Kriyol
Sãotomense	43% (3/7) ⁴	43% (3/7)
Principense	25% (2/8)	13% (1/8)
Annobonese	13% (1/8)	25% (2/8)
Angolar	25% (2/8)	25% (2/8) ⁵

By way of contrast, note in Table 4 how high a percentage of wh-words the Gulf of Guinea creoles have in common:

Table 4. Percentage of wh-words shared among the Gulf of Guinea Portuguese creoles⁶

	Sãotomense	Principense	Annobonese
Principense	71% (5/7)		
Annobonese	63% (5/8)	50% (4/8)	
Angolar	71% (5/7)	50% (4/8)	63% (5/8)

Importantly, the differential retention of Portuguese wh-words can be accounted for under an analysis which places the emergence of these question words at the genesis of the creoles, since the particular contexts in question offer explanations for the discrepancy. As Ferraz (1983) notes, there was a relatively large Portuguese population relative to the presence of Africans in Cape Verde and in what today is Guiné-Bissau, whereas when Sãotomense,

Principense, Annobonese and Angolar (the Gulf of Guinea creoles) formed, the proportion of Portuguese to Africans was considerably less (see also Thomason & Kaufman 1988:156). In the early 16th century, due to disturbances on the plantations, São Tomé entered into a steep decline leading to the exodus of large numbers of Portuguese settlers to Brazil and elsewhere (Ferraz 1983:123). Ferraz notes that, “in linguistic terms, the departure of the Portuguese meant the removal to some extent of the Portuguese base from the situation of language contact”. This, in turn, led to pronounced influence from the African substratum, not only in São Tomé, but in Príncipe and Annobon as well, given that these islands were dependent on São Tomé for inhabitants. African influence was particularly profound upon Angolar, which developed on São Tomé among escaped slaves in isolation from Portuguese speakers. Thus the differing degrees of distance from Portuguese between these two groups of creoles is traceable to early demographic ratios — and thus to the period when Africans were in a position to accommodate to Portuguese in creating a wh-word inventory.

2.1.2. *Wh-word constructions*

The use of wh-words in sentences is much more uniform across the Portuguese creoles of West Africa. The pattern found in all these creoles except Annobonese is:

WH-WORD + [kV] + SENTENCE

Kabuverdianu:

- (2) a. **Kònte** **k'** **bo** **kre?**
 how.much FOC 2SG want
 ‘How much do you want?’
 b. **Ken ki** **ta** **fla** **ma** **bu** **ten** **48** **anu?**
 who FOC PROG say COMP 2SG have 48 year
 ‘Who says that you are 48 years old?’ (Cardoso 1989:34-35)

Kriyol:

- (3) a. **Kin ku** **fala** **u** **sin?**
 who FOC say 2SG so
 ‘Who told you so?’
 b. **Ke** **ku** **bo** **misti?**
 what FOC 2PL want
 ‘What do you want?’ (Kihm 1994:230-31)

Principense:

- (4) a. *Kwá ki ci té â?*
 what FOC 2SG haveQU
 ‘What do you have?’
 b. *Kwe mādá ki ti pagá ufogu ã?*
 why FOC ANT put.out fire QU
 ‘Why did you put out the fire?’
 (Günther 1973:93-94, Ferraz & Valkhoff 1975:25)

Angolar:

- (5) a. *Kwai ma ene zi?*
 what FOC 3PL do
 ‘What did they do?’
 b. *Dia kutxi ma bô ka ba kwanda?*
 day which FOC 2SG FUT go up
 ‘What day are you going up?’ (Maurer 1995: 137-8)

Sãotomense:

- (6) a. *Keñé ku bi aí kε nō òti.*
 who FOC come here house our yesterday
 ‘Who came here to our house yesterday?’
 b. *K(w)ātu boy k(u) ina te ni lósa?*
 how.many ox FOC 3PL have in land
 ‘How many oxen do they have on their land?’
 (Ferraz and Valkhoff 1975:25-6)

Annobonese:

- (7) a. *Xa bo fala?*
 thing 2SG speak
 ‘What do you say?’ (Post 1995:199)
 b. *Quengue bo sa ja jua?*
 who 2SG PROG look.for
 ‘Who are you looking for?’ (Barrena 1957:42)

With the exception of Annobonese which contains no focus or interrogative particle, the uniformity exhibited by these creoles can be accounted for if it is assumed that in the creolization process the Africans adapted the structure of their languages to that of Portuguese. Many relevant languages of the African west coast have an analogous element in their wh-construction, though it is not

always a focus particle *per se*. We cite here examples from several West African languages: Wolof (relevant for Kriyol and Kabuverdianu), and Yoruba, the Akan varieties Ashanti and Twi and the Gbe varieties Ewe and Gen (relevant for the Gulf of Guinea creoles):

Wolof:

- (8) **L-an** *nga-y def?*
 what-INT 2SG doing
 ‘What are you doing?’ (Gamble 1991:58)

Yoruba:

- (9) **Kíni ni** *mo rà?*
 what FOC I buy
 ‘What am I buying?’ (Boretzky 1983:225)

Ewe:

- (10) **Núkà nè** *dà égbè?*
 what INT cook today
 ‘What have you cooked today?’ (Pasch 1995:80)

Gen:

- (11) **Nú-ke** *mù wɔ ná wó ò?*
 what-INT 1SG do PREP 2SG INT
 ‘What have I done to you?’ (Pasch 1995:79)

Ashanti:

- (12) **Àkyíná édién nà** *wèú-ú-yè?*
 tomorrow what FOC 2SG-do-PROG
 ‘What are you doing tomorrow?’ (Boakye 1982:328)

Twi:

- (13) **Deen na** *wo-be-di?*
 what FOC eat-FUT-2SG
 ‘What will you eat?’ (Adu-Amankwah and Botne 1996:44)

In the 15th century, as today, Portuguese had a focus construction (14a), and a corresponding interrogative focused construction (14b). This construction soon generalized to become the default question formation structure in Portuguese:

Middle Portuguese:

- (14) a. *É por esto **que** vos digo.*
 is because.of this that you.OBJ say.1SG
 'It's because of this that I tell you.' (Huber 1933:285)
- b. *Quem é a **que** parir-a?*
 Who is she that give.birth-3SG.FUT
 'Who is the one who will give birth?'
 (Carvalho Buescu 1984:143-4)

It is highly probable that the general pattern found in the West African Portuguese-based creoles is the result of mutual linguistic accommodation. In all the creoles in question except Angolar, the focus element is *k(V)*, which corresponds to the Portuguese focus element *que*.

Ma in Angolar (< *kuma* 'what') (< Middle Portuguese *coma* 'how' < colloquial Latin *quomodo* + *ac*) is also used as a focus element in declarative sentences:

Angolar

- (15) *Maya thô **ma** no bê.*
 Maya EMPH FOC 1PL ver
 'It is Mary we saw.' (Maurer 1995:136)

Interestingly, Angolar was created by slaves that had escaped from a São Tomé plantation around 1517, at which time Ferraz (1974) believes that their plantation variety of Portuguese had not completely stabilized. They lived isolated in the interior of the island until the latter part of the 19th century. The fact that their language has the focus element *ma* rather than *ki* suggests that their creole did in fact form independently, at least in part, from the other creoles.

The reason for the absence of an interrogative particle in Annobonese question formation is unclear. Mutual linguistic accommodation would predict its presence, yet we do not find it, although it may well be that such a particle is used optionally.

Table 5 shows that the principles of mutual linguistic accommodation correlate quite well with attested constructions in these creoles:

Table 5. Predictions from mutual linguistic accomodation principles and attestations in West African Portuguese creoles (predictions [P] and attested findings [Att.] are indicated in terms of presence [+] and absence [-]):

	KV	Kr	ST	Pr	Ang	Ann
Focus	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: -
Fronting	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +
<i>In situ</i>	P: - Att: -	P: - Att: -	P: - Att: -	P: - Att: -	P: - Att: -	P: - Att: -

2.2. Asian Portuguese Creoles

The respective contact situations which gave birth to the Portuguese creoles in Daman and Korlai are relatively simple in that both emerged as the result of contact between primarily two languages: Gujarati and Portuguese for Daman Portuguese (DP) and Marathi and Portuguese for Korlai Portuguese (KP). Both Marathi and Gujarati are Nearyan languages with SOV order. Kristang (Kri) formed from Portuguese and Malay (SVO), with possible influence from Hokkien Chinese (cf. Hancock 1975); this would have been a three-language contact situation, but the influence of Hokkien Chinese on Kristang was weaker than that of Malay. Sri Lankan Portuguese (SLP) formed in a two-language contact situation between Portuguese and Tamil (SOV); the Dutch took over the island later but this affected the language only in its lexicon.

2.2.1. Wh-word inventories

As is evident in Table 6, Portuguese wh-words form the basis of all of Daman Portuguese's wh-words (8/8), whereas Korlai Portuguese has 6/8 (75%), Kristang 5/8 (63%), and Sri Lankan Portuguese 4/8 (50%):

One plausible reason for the large percentage of direct transfer from Portuguese to DP and KP is the nature of the respective contact situations. Originally, there was no intention on the part of the Portuguese to even settle in India permanently. Rather, they initially had intended to use Indian ports as temporary bases for commerce. They did end up settling there in the end, but their sole objective was commercial, not agricultural: at no time did the Portuguese have the intention of developing plantations in India. Thus, the ratio of slaves to Portuguese soldiers was relatively low, a situation which lent itself to higher access to the variety of Portuguese used to communicate with them.

Table 6. Comparison of *wh*-words in Asian Portuguese-based creoles

Ptg	<i>quem</i> 'who'	<i>que</i> 'what'	<i>quando</i> 'when'	<i>onde</i> 'where'	<i>porque</i> 'why'	<i>como</i> 'how'	<i>quanto</i> 'how many'	<i>qual</i> 'which'
DP	kě	kik ki koiz	kwan, kɔm, kyɔr	(d)un, (19th c. ond)	parki	kilay, kom	kwant	kal
KP	kē	ki	kɔr, kɔr ki, ki	un, kalwan	pəri (<pərki) keku	kile	kāt	kal
Kri	keng	ki	kora, ki tempu	undi, ki banda	kifoi, ki kausa	klai	kantu, kai tantu	kal ńgua
SLP	keem	kii	kii ɔɔra	oondi	kii-pa, kii viida (porkii, parkii)	kilaay	kaantu	kii

(Sources: research by the author, Baxter 1988, Ian Smith, p.c.)

There are, however, three bi-morphemic *wh*-words in KP and DP:

- a. KP *kile*, DP *kilay* from *que laia* 'what manner'
(cf. KV *modi*, Gulf of Guinea *mɔd(i)* from *que modo* 'what manner')
- b. KP *kɔr* and DP *kyɔr* from *que hora* 'what time'
(cf. Kr *kalhora*, Gulf of Guinea *dia kutxi*, *ola kutxi*, , etc. from *qual hora/*
dia 'which time / day' (*kutxi* < Kimbundu 'which' [Da Silva Maia
1964:55-6]), *ku ola* from *que hora* 'which day')
- c. KP *kalwan* from *qual banda* 'which side'

Bi-morphemic question words can emerge in a creole as an expression of universal tendencies towards semantic transparency. However, this does not seem to have been the source of the bi-morphemic *wh*-words in these creoles for the most part.

KP *kɔr* may have formed due to influence from Africa, but may also have formed independently; KP *kalwan*, however, seems traceable to substrate influence, specifically calquing upon Marathi's bimorphemic *wh*-word for 'where', *kay badzu-la* (which side-DIRECTIONAL) 'where'.

DP's monomorphemic alternate for 'how' *kom*, from Portuguese *como*, is

an acrolectal latter-day incorporation. The bimorphemically derived *kilay* was still in use at the end of the 19th and the beginning of the 20th century, as attested in Dalgado (1903), but is now heard predominately in the speech of the oldest speakers, being replaced gradually by *kom* as the result of decreolization. Daman was part of Portugal until 1961, since which time the Portuguese language has continued to maintain a presence in schools and the church. Meanwhile, in Kristang there are many bi-morphemic wh-words, most due apparently to the fact that Malay has bi-morphemic question words (Baxter 1988:189); Hokkien Chinese may also have played a role here. However, *klai* 'how' and *kora* 'when' are most likely due to influence from Indo-Portuguese varieties.

The use of *kii* for 'which' in SLP is due to generalization, while the departure from Portuguese evidenced by *kii-pa* is due to the widespread use on SLP grammar of *-pa* (< *para* 'for') as an oblique marker.

2.2.2. Wh-word constructions

DP, KP and Kristang coexist with Gujarati, Marathi and Malay and Hokkien respectively. As SOV languages, Marathi and Gujarati have the expected *in situ* wh-constructions as in the Marathi examples in (16):

Marathi:

- (16) a. **Kon** *ala*?
 who came
 'Who came?'
 b. *Anjana kay sangitle*?
 Anjana what said
 'What did Anjana say?'
 c. *Anjana puskək kon-la dile*.
 Anjana book who-OBJ gave
 'Who did Anjana give the book to?'

This structure does not exist in Portuguese. Because of substantial typological distance between Marathi/Gujarati and Portuguese, mutual linguistic accommodation would not come into play. Instead, it is reasonable to predict that Marathi and Gujarati speakers shifting to Portuguese would carry this structure with them into the language they were attempting to learn. In contact situations in which there is substantial typological distance between languages, shifting speakers will often rely on what for them is most unmarked,

most universal. Given that Marathi and Gujarati have *in situ* question formation while Portuguese has fronting, the Marathi and Gujarati speakers shifting to some version of Portuguese would do what was for them most economical, which would be to leave the question word in its place instead of learning a new construction.

This prediction is correct. The fronting option found upon elicitation in DP is an acrolectal trait of the language, without a doubt not the default option, as the author's work has ascertained:

Daman Portuguese:

- (17) a. *Minh boc qui-lai ha chega alli no ramad?*
 my mouth how FUT arrive there to branch
 'How will my mouth get to the [grapes on the] branch?'
 b. *Oss un' já foi?*
 2SG where PAST went
 'Where did you go?'
 c. *Oss ki já troç?*
 2SG what PAST brought
 'What did you bring?' (Dalgado 1903:17, 21, 22)

Korlai Portuguese:

- (18) a. *Buk pɔɔ kɛ yadew?*
 book 2SG.OBJ who gave
 'Who gave you the book?'
 b. *Use oj mused un ti andad?*
 2SG today morning where had gone
 'Where did you go [and come back from] this morning'
 (Clements 1996)

In reference to Kristang, its substrate language Malay allows *in situ* and preverbal placement but not fronting:

Malay:

- (19) a. *Dia tengak apa?*
 he see what
 b. *Dia apa tengak?*
 he what see
 'What does he see?' (Alan Baxter, p.c.)

Portuguese, however, allows only fronting. In Kristang, mutual linguistic accommodation did apply in that, of the three options that the two contact languages allowed, all three are found (with the proviso that subject-verb inversion is ungrammatical):

Kristang:

- (20) a. *Úndi bos ja parí?*
 where 2SG ANT born
 ‘Where were you born?’
 b. *Bos úndi ja parí?*
 2SG where ANT born
 ‘Where were you born?’
 c. *Eli ta bai úndi?*
 he PROG go where
 ‘where is he going?’ (However, **Úndi ta bai eli?*)
 (Baxter 1988:189, p.c.)

The situation in Sri Lankan Portuguese is similar. Wh-word placement is quite free, but as in DP and KP, there is a tendency for wh-words, particularly questioned arguments, to appear immediately before the verb (in situ):

Sri Lankan Portuguese:

- (21) a. *Avara boos kii kera faya?*
 now 2SG.NOM what INT do
 ‘Now what are you going to do?’
 b. *Avara kii kera, boos-pa?*
 now what want you-DAT
 ‘Now what do you want?’
 c. *Aka kaantu ɔɔra-pa pooy ovii, eev?*
 that how.many hour-DAT can hear 1SG-NOM
 ‘At what time can I hear it?’
 d. *Avara kaantu aanu avara akii jaa-vii, botus?*
 now how.many year now here PAST-come you.HON
 ‘How many years (has it been) since you came here?’
 (Ian Smith, p.c.)

This type of structure is foreign to Portuguese and is thus attributable to Tamil. Thus mutual linguistic accommodation via substrate influence applies, as this structure was introduced by speakers shifting from Tamil who found it easier

to use a structure which did not involve movement than take as a default rule one which involved fronting. The shifting Tamil speakers can also be argued to have chosen on the basis of universal tendencies, given that in-situ question formation involves no movement, whereas fronting does involve movement, or at least indexing.

Table 7 summarizes the accuracy of predictions as to whether the creators of these creoles would utilize mutual linguistic accomodation or universal tendencies, in developing wh-word constructions:

Table 7. Predictions as to source of wh-word constructions in Asian Portuguese creoles (predictions [P] and attested findings [Att.] are indicated in terms of presence [+] and absence [-] of universal tendencies [UT] or mutual linguistic accomodation [M]):

	KP	DP	Kri	SLP
Preverbal	N/A	N/A	P: M + Att: +	N/A
Fronting	P: UT - Att: -	P: UT - Att: - (only acro.)	P: M + Att: +	P: UT - Att: -
In situ	P: UT + Att: +	P: UT + Att: +	P: M + Att: +	P: UT + Att: +

3. Wh-words and Wh-constructions in French-based Creoles

3.1. Wh-word inventories

Table 8 is a collection of the wh-words in seven French-based creoles. Valdman (1978:207) notes that “in general, interrogative pronouns and adverbs are formed by adjoining the particle *ki* to an indefinite noun: *moun* ‘person’; *sa, bagay* ‘thing’; *lè* ‘hour’; *tan, kan* ‘time’; *jan, manière* ‘manner’; *koté (kòt), bò, trou, kartié* ‘place’. These nouns are spread very widely but vary from one variety to another”. The forms in bold are the ones considered to be derived from the corresponding French wh-words.

Regarding which items these creoles share with French⁷, note that St. Lucian, Haitian and Guadeloupean cluster on the low end of the scale (50%-57%), Tayo is at the top (100%), and the Indian Ocean creoles are in the middle (63%-88%).

Table 8. Comparison of *wh*-words in French-based creoles

Fr	qui	que, quoi	quand	ou	pour- quoi	com- ment	com- bien	quel, qui, quoi	de qui
M	(sanla) ki, ki	((sanla)) ki, ki ,	kah , ki ler	(ki) kote	kifer	ki manyer, kuma	ko- myeh	ki	
S	ki , ki sen- la	ki , ki sa, ki si-sa, ki sa-sa	kâ , ki ler, kel- er	kot(e) ki ladrua, oli	akoz (ki), ki fer, ak- fer	koma , ki maÿer ki fasô	kôbiê	lekel	pur ki
T	se ki , (ke si ka, ke se ki)	(se) kwa	ka	u	pukwa de kwa	(se) koma	kobja	(le) kel	
SL	(ki) mun	(ki) sa, ki	ki lè ki tã	(ki) kote, ko	pou- tchi	ki maniè, kumã	kômê	ki lès	
H	ki moun	(ki) sa	ki lè	(ki) kote, ki bò	pou- kisa	ki jan, kou- man	kouben	ki lès	
G	kimoun	kisa kibiten	kitan, kilè	ola , kikote	pou- kisa	ki jan, kiman- nyé	ko- myen	kibèt	
L	ki , sa-k(i), ki-k(i), ki-se- (sa)- k(i)	ki , ki sa, ki-se- (ki), sa-ki	ekô , ke lœr	eu , au	kofe	komô	kôbjě	ke(l)	pu ki

(M = Mauritian Creole, S = Seychellois Creole, T = Tayo Creole, SL = St. Lucian Creole, H = Haitian Creole, G = Guadeloupean Creole, L = Louisiana Creole; Sources: Baker 1972, Corne 1997a, Erhardt 1993, Carrington 1994, Albert Valdman, p.c., Cérol 1991, Neumann 1985)

All seven have a reflex of French *qui* ‘who’ and *combien* ‘how many’; all but Tayo share a reflex of *qui*, *quand*, and *combien* (38% of the French items), and all but Tayo and Louisiana Creole share a reflex of *qui*, *quand*, *comment*, and *combien* (63% of the French items). Interesting is that the Caribbean creoles (Haitian, St. Lucian, and Guadeloupean) share six of eight *wh*-words

(St. Lucian and Haitian all eight), while the Indian Ocean creoles Seychellois and Mauritian have seven of eight in common, suggesting very close relationships. Based on a detailed comparative study of morphosyntactic phenomena, Bollée 1977a concludes that the latter two are the most closely related of the Indian Ocean creoles. Revealing is that this conclusion is possible based solely on *wh*-words, which is positive evidence for the claim that *wh*-words can often constitute a reliable diagnostic mechanism for determining relationships between creoles.

3.2. *Wh-constructions in the French-based creoles*

In contrast to the Portuguese-based creoles in Africa, which homogeneously have *ki* or its equivalent as a question particle, the picture is variable within the French-based creoles. In French, one finds both fronting and *in situ* formation of the type *Tu as vu quoi?* (lit. 'you have seen what'). In the French-based creoles under consideration, Tayo prefers *wh*-words *in situ* but allows *wh*-word fronting, while Seychellois appears not to have the option of *in situ* *wh*-word placement, but all others appear to use *wh*-words in both fronting and *in situ* constructions, with fronting as the default option. These outcomes can be attributed to mutual linguistic accommodation.

3.2.1. *Indian Ocean creoles*

Mauritius was settled by 1715, and by 1723, there were 182 inhabitants, of which 6 (3%) were French-speaking Europeans from Réunion, 112 (62%) French speakers directly from France, and 64 (35%) Malagasy-speaking slaves from Madagascar. Between 1727 and 1730 the slave population grew from around 30 Malagasies to about 1000; approximately 600 were West Africans speaking mainly Wolof, Fon and Bambara, while the rest were from India (speaking Tamil and Bengali) and Madagascar (speaking Malagasy). Around 1730, there were also 400-600 French-speaking settlers not having come via Réunion. Baker & Corne (1982:169) argue that a creole developed out of this situation and was stable by the 1770s. The major sources of slaves between 1735 and 1770 were from Madagascar, and then from East Africa. Baker & Corne (1982:176) identify four languages from the relevant area which were important: Makhuwa, Yao, Mwera and Bemba, all Bantu languages.

Based on the situation which existed at the beginning of the inhabitation of Mauritius between 1720 and 1740, we can assume that the structure of

Mauritian would have been influenced by French, West African languages, the aforementioned East African languages, and Malagasy. From the standpoint of shared structures, French and Malagasy share *in situ* question formation, while French and the African languages share wh-word fronting. Moreover, French has a cleft construction which is structurally comparable to a focus construction in East African Bantu languages (cf. [8] - [13]), as in Bemba:

Bemba:

(22) *N-áni úmu-káshána áapéela úmu-pila?*

FOC-whothe-girl give CL-ball

'Who does the girl give the ball to?' (Oger 1982:41)

Thus we would expect that Mauritian would have both wh-word fronting and *in situ* question formation, and that it show signs of question particle use. Indeed, this creole does allow both of types of question formation (Baker 1972), and furthermore does have a question particle *ki*, although it is used optionally (Anand Syea, p.c.). The notion of mutual linguistic accommodation makes the correct predictions in this case.

The Seychelles islands were claimed by France in 1756, with the largest, Mahé, inhabited in 1772. After 1788, Mahé's population increased substantially: from 250 in 1788 (221 slaves, 20 whites [8%], 9 freedmen), to 6963 in 1825 (6058 slaves, 582 whites [8%], 323 freedmen) (Bollée 1977:3, Holm 1989:401). The islands came under British rule in 1810, though English never replaced the creole. Slavery on the Seychelles was abolished in 1835 and over the next 40 years the islands became the locus of release for many freed captives. Thus, during approximately the first 100 years of occupation (i.e. between 1772 and 1885), the percentage of whites to slaves hovered between 8% and 12%. The slaves and freedmen were African, most probably East Africans speaking Bantu languages (cf. Baker & Corne 1982:100-101). As we have seen, the relevant Bantu languages often use a focus particle in question formation. Thus in Seychellois we would expect to find a question particle given the presence of Bantu-speaking Africans and the presence of a comparable structure in French. Moreover, we would expect wh-word fronting since Bantu languages and French front the wh-word. We would not, however, expect *in situ* question formation because it is only found in French and the presence of French speakers was only between 8% and 12% during the first 100 years of the creole's formation.

These predictions are borne out by the data available to us. In Corne

(1977:187-9) and Bollée (1977b:779-81), we find only wh-word fronting and no *in situ* question formation (though we suspect the latter may be possible). *Ki* is indeed used as a focus particle (optionally in most cases; for some informants seemingly obligatory in questions involving ‘who, what’; not with *kâ* ‘when’ and *koma* ‘how’, rarely with *kot(e)* ‘where’) (Corne 1977:187-88).

3.2.2. Caribbean French creoles

Haiti was first settled in 1631. Buccaneers were encouraged to settle there in the second half of the 17th century, and by 1697 Haiti was in the process of becoming a classic sugar colony. During the early period of the settlement, most of the slaves came from other European colonies. Later, however, the French established their own slave trade. The slave population increased from 2,000 in 1681 (33% of the total population) to 165,000 in 1753 (91% of the total population). Between 1711 and 1740, most of the slaves working on the plantations were speakers of Mande (in particular Bambara, Malinke, and Dyula) and Kwa (Ewe-Fon) (Holm 1989:382-384). While the Kwa languages display a question particle and wh-fronting, Bambara (cf. Long 1970 and Gospel Missionary Union 1964) apparently does not have a particle but exhibits wh-fronting as well as *in situ* question formation. We find the question particle *ki* in Haitian (though only optionally, when the subject is questioned as in *Komben moun (ki) vini?* [lit. ‘how-many people [FOC] come’] ‘How many people came?’ [A. Valdman, p.c.]). Both *in situ* placement and fronting are found, as would be expected given what we find in Bambara, Ewe, and Fon.

St. Lucia was first settled by Frenchmen in 1650. It changed hands between the British and French many times, until in 1763 it became French, with a considerable influx of French planters and their African slaves ensuing (Holm 1989:374-75). Given this background, we would expect the question particle *ki*, which is the case, though its use is limited and, in this, strikingly comparable to that of Haitian. St. Luc allows both *in situ* placement and fronting, again predicted because French and West African languages display both wh-fronting and *in situ* question formation.

Finally, the first settlers arrived on Guadeloupe in 1635 and in the late 1600s, Guadeloupean (and other creoles in the Lesser Antilles as well) stabilized, with fundamental restructuring due to the influence of African languages (Holm 1989:365). Given this background, we again would expect *ki* as a question particle, but do not find this attested in Cérol (1991). Further

research would be needed to verify this. Moreover, we would expect the *in situ* option in question formation, but no such examples are found in Cérol (1991), though we suspect it may be an option.

3.2.3. *Louisiana Creole*

Louisiana Creole was formed in a situation of roughly equal numbers of slaves and settlers between 1700 and 1730-1740. A pidgin French similar to the modern creole is attested as early as 1730 (Holm 1989:387-88). It may be due to the nature of the contact situation in Louisiana that the question particle *ki* is not found there despite its prediction on the basis of the principles of mutual linguistic accommodation. The development of the question particle *ki* may have been precluded due to the fact that question words in the creole contain *ki* in final position, e.g. *sa-ki*, *kel ki*, etc. (Neumann 1985:333-338).

3.2.4. *Tayo*

Tayo developed on the island of New Caledonia, where there was never a plantation culture, nor mass influxes of laborers. New Caledonia was annexed by the French in 1853. The creole spoken there emerged from the contact among speakers of several varieties of French, speakers of Melanesian languages, and speakers of Chinese, Indonesian, and Japanese (Ehrhart 1993:7). There is clearly a strong French influence in its formation, given, for example, that all of its question words are directly taken from French (i.e. there are no bimorphemic *wh*-words). In any case, under these conditions we would not expect to find *ki* as a question particle, and would expect to find both fronting and *in situ* placement. Accordingly, we find nothing of the first and extensive evidence of the second and third (cf. Ehrhart 1993:196-98).

Table 9. Predictions from mutual linguistic accomodation principles and attestations in French-based creoles (predictions [P] and attested findings [Att.] are indicated in terms of presence [+] and absence [-]):

	M	S	H	SL	G	L	T
Focus	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + At: -	P: + Att: -	P: - Att: -
Fronting	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +
<i>In situ</i>	P: + Att: +	P: - Att: -	P: + Att: +	P: - Att: - (?)	P: - Att: - (?)	P: (+) Att: +	P: + Att: +

To conclude this section, we give this tabular summary of the findings for the French creoles discussed.

In the case of the question particle *ki*, mutual linguistic accommodation makes the correct prediction five of seven times (75%). For fronting, the predictions are borne out 100% of the time, but it must be said that, as was pointed out in section 1 above, cross-linguistically fronting is extremely common. Thus, these particular predictions here are not particularly meaningful.

The predictions involving *in situ* placement of *wh*-words are more interesting. The option of *in situ* placement is to be expected in Louisiana Creole due to influence from French alone, despite its absence in languages of Senegal like Wolof, since it emerged in a situation of roughly equal numbers of slaves, mostly from Senegal, and white French speakers, such that the slaves would have had direct and frequent contact with French speakers. This special case is marked in the table by a parenthesis around the “+”. For Tayo, as the Melanesian languages of New Caledonia are VOS/SVO and allow both fronting and *in situ* placement, one would expect the same options in the creoles, which are indeed found. The wrong predictions in the cases of St. Lucian and Guadeloupean are based on Carrington (1984) and Cérol (1991) respectively, where *in situ* placement was neither mentioned nor documented. We assume, however, that the option probably exists and would be revealed by further research.

So far, we have been able to show for the Portuguese- and French-based creoles examined that in most cases, the principles of mutual linguistic accommodation account well for the types of question formation they exhibit. We have also seen that the working hypothesis that the more *wh*-words two or more creoles have in common, the more closely they are related to one another seems to be tenable so far.

In the subsequent sections, we will examine *wh*-words and question formation in selected Spanish- and English-based pidgins and creoles to see to what extent our results from sections 2 and 3 are applicable to these languages as well. What we will find is that, while mutual linguistic accommodation continues to account for the nature of question formation, shared *wh*-words

are not necessarily an indication of direct relatedness during pidgin and creole formation.

4. Wh-words and Wh-constructions in Two Spanish-based Creoles

4.1. Wh-word inventories

The Spanish-based creoles Papiamentu and Palenquero share five forms from

Table 10. Comparison of wh-words in Papiamentu and Palenquero

Span	<i>quién</i> 'who'	<i>qué</i> 'what'	<i>cuándo</i> 'when'	<i>dónde</i> 'where'	<i>por qué</i> 'why'	<i>cómo</i> 'how'	<i>cuánto</i> 'how many'	<i>cuál</i> 'which'
Pap	kende ken	ki, kiko	ki ora, ki dai, ki	unda, na unda tempu	dikon, pa(si)- kiko	kong	kuinto	kua, ki
Pal	kien, kien	ke kusa, ke	a ké ora, ki ria, kuando	andi, (a)onde	pogké, kotá	kumu	kuinto	kuale

(Pap: Bartens 1995:252-53, Philippe Maurer [p.c.]; Pal: Armin Schwegler [p.c.]

Spanish, the equivalents of ‘who’, ‘what’, ‘how’, ‘how much/many’, and ‘which’, and seven forms (some departing from Spanish) with each other. Papiamentu shares 63% (5/8) of its question words with Spanish, Palenquero 88% (7/8).

Of the bimorphemic wh-words exhibited by these creoles (Papiamentu *hende* < *qué gente*, *kiko* < *qué cosa*, *ki ora/dia/tempu*, *dikon* < *de cómo*, *pa(si)kiko* < *para [hacer] qué cosa*; Palenquero *ke kusa*, *ke ora/ria*), it is interesting to note that they share the equivalents of both ‘what’ (*kikos* and *ke kusa*), ‘when’ (*ki ora*, *ki dia* and *a ke ora*, *ki ria*) and ‘where’ (Papiamentu *andi*, *(a)onde* and Palenquero *unde* (from Portuguese *onde*). Thus, rather than common origin these commonalities reveal three points of origin: 1) Spanish, 2) Portuguese (‘where’), and in the case of the bimorphemic forms, 3) either the tendency toward semantic transparency as part of universal tendencies in

pidgin/creole formation and/or African substrate influence. We will comment further on this below.

5.2. *Wh-word constructions*

Pap fronts the *wh*-word with no other alteration such as the addition of a question particle or subject-verb inversion as found in Spanish. According to Philippe Mauer (p.c.), Papiamentu does not display *in situ* question formation:

- (23) a. **Kiko** *bo ta bende?*
 what 2SG PROG sell
 ‘What are you selling?’
 b. **Ken** *bo ta kere ku ta parse mi tata?*
 who 2SG PRES believe that PRES resemble my father
 ‘Who do you believe resembles my father?’
 (Kouwenberg & Muysken 1995:216)

The situation is identical in Palenquero: Friedemann & Patiño (1983:168-70) report that after *wh*-word fronting, there is no inversion, nor is there any trace of a question particle similar to *ki* in the Portuguese-based creoles in Africa. Thus, question formation in both creoles consists simply of a fronted *wh*-word. How do the structures reflect the respective creoles’ sociohistorical background? We will now attempt an answer to this question starting with Papiamentu.

Currently, Papiamentu is spoken on the leeward Netherlands Antilles, Curaçao and Bonaire, as well as the island of Aruba.⁸ Spanish explorers first encountered Curaçao in 1499, and the other two islands shortly thereafter. In 1527, the Spanish crown established a small garrison of around 25 soldiers on Curaçao. Up until 1634 the Spanish shared the island with the original inhabitants, Arawakan-speaking Amerindians. In this year, the Dutch took over Curaçao, forcing the Spaniards and most of the indigenous inhabitants to the mainland. Several hundred Dutch settled Curaçao, and in 1688, the Dutch also seized Aruba and Bonaire. The Dutch often understood Spanish because of this nation’s political affiliation with the Spanish empire until 1581. Moreover, they frequently knew Portuguese and used Portuguese-based creoles because of their seizure of what had been up to that time Portuguese colonies. It is important to note that the Dutch reserved the use of Dutch for speaking among themselves, and did not practice miscegenation as did the Portuguese.

The first slaves began to arrive in Curaçao in 1650. The highpoint of the slave trade was the last third of the 17th century, when Curaçao became the slave Dutch entrepôt for the whole of the Americas. In Curaçao, in 1683, 75% of the slaves permanently living there worked as house slaves, while the other 25% were engaged in plantation work.

Another important group instrumental in Papiamentu's formation and development was the community of Sephardic Jewish plantation owners. These had left Spain and Portugal in the 16th century and had established plantations in the north of Brazil. Around 1658, Portugal began to retake the northeastern Brazilian coast from the Dutch, prompting an exodus, with Dutch plantation owners, from Brazil into Dutch colonies. Both the Sephardic Jews and the Brazilian Dutch brought slaves with them and these undoubtedly spoke a variety of pidgin Portuguese. Moreover, the Jewish plantation owners spoke one or more of a number of Iberian-based languages — Portuguese, Spanish, Judeo-Portuguese, Judeo-Spanish, Galician or other regional varieties. The Brazilian Jews came to play a principal role in the slave trade on the island, the Spanish granting them a monopoly in supplying their American colonies with slaves. These former plantation owners most likely extended the restructured Portuguese they had been using with their slaves to use with the new slaves. Furthermore, in all probability, at least some of these new slaves knew pidgin Portuguese from their contact with it on the African west coast. Important in Papiamentu's formation was that the Sephardic Jews had lower social status among whites than the Dutch, and thus had more contact with the slaves.

According to Bartens (1995:247), the formative period for Papiamentu seems to have been between 1650 and 1700. In this case, crucial input would have come, on the one hand, from the various speech varieties of the Sephardic Jews and, on the other, from the slaves, some of whom had probably already learned a Portuguese-based pidgin in Brazil or Africa.

From the perspective of mutual linguistic accommodation, we might expect less influence from African languages because Papiamentu formed in South America with significant input from Europeans (the Sephardic Jews), as well as with input from two distinct groups of Africans, one already conversant in pidgin Portuguese, and the other only recently exposed to it. The one feature in question formation which all languages in the contact situation had in common was fronting of the *wh*-word, and it is precisely this feature that we find in Papiamentu. The heterogeneity of the linguistic situation in which the

creole formed renders it unsurprising that it does not share all of its question words with Spanish: the presence of the bimorphemic items suggests that universal tendencies, substrate influence or simply relexification of pidgin Portuguese forms may have played a role.

Palenquero⁹ is spoken almost exclusively by the older members of El Palenque de San Basilio, a village of around 4000 inhabitants south of Cartagena, Colombia. The creole originated in the early 1600’s when around 30 escaped slaves built a fortified village in the interior. Palenquero may have developed out of a more widespread Spanish-based creole that may have existed at one point in the New World. Extralinguistic evidence (e.g. certain magic rituals) suggests that many of the slaves who founded El Palenque de San Basilio were from Kikongo- and Kimbundu-speaking areas, leading Granda to propose that Palenquero has a genetic relationship with Sãotomense and Annobonese (cf. Bartens 1995:268-69).

With respect to question formation, Palenquero displays simple *wh*-word fronting with no inversion or addition of a question particle. Although Spanish does have a focus construction with *que* (¿*Qué es lo que queremos?* ‘What is it that we want?’), it is used only with *qué*, and in general not as frequently as in Portuguese and French. Meanwhile, Kikongo does not have a focus particle. Thus, it is to be expected that Palenquero lacks one.

Table 11. Predictions as to source of *wh*-word constructions in Spanish-based creoles (predictions [P] and attested findings [Att.] are indicated in terms of presence [+] and absence [-]):

	Pap	Pal
Focus	P: - Att: -	P: - Att: -
<i>In situ</i>	P: - Att: -	P: - Att: -
Fronting	P: + Att: +	P: + Att: +

5. Wh-words and Wh-constructions in English-based Creoles

5.1. *Wh*-word inventories

Table 12 compares *wh*-words in a representative range of English-based pidgins and creoles. Where possible, we have also included *wh*-words found

Table 12. Comparison of *wh*-words in English-based pidgins and creoles

	<i>who</i>	<i>what</i>	<i>when</i>	<i>where</i>	<i>why</i>	<i>how</i>	<i>how much/ many</i>	<i>which</i>	% shared with Eng.
J	huu (-dat)	wa(t) , we, wara (rural)	wen- taym, wen	we , we- paat, wich- paat	wa- mek, (h)ou- kom	ou	(ho)m- och	wich	88% (7/8)
Sr	suma (<i>oe som- body</i>)	san	oten (<i>oe tem</i>)	pe (<i>oe (plesse)</i>)	fu san ede (<i>ver wate</i>)	fa (<i>oe fasse</i>)	o- meni	sortu	13% (1/8)
D	suma	san	on ten	pe	fu saide meke (fu sai)	fa	on men	sootu, ondi	13% (1/8)
Sm	Ambé (<i>hudi</i>)	andí	(na)- ún-té, (na)- ún-júu, (<i>hutem</i>)	un- kamía, (na)- un-sé, naasé (<i>husei, hu- gamja</i>)	fu andí édi, fu andí mbéi	ún-fa, fa, un (<i>hufa</i>)	ún- meni	ún, ún-di (<i>hu-</i>)	0% (0/7)
K	u , udat	weytin	ustem, wey	wey , usay, us(pat)	weytin du, weytin mek	aw	omos	us	63% (5/8)
C	wis- man, hu , hus- man, hus	wa ting, weti, hus- kayn tin	hu- taym, wisty	wusai, wisay	fo sek we- ting, way	ha	ha- meni	wus- kayn	50% (4/8)
TP	husat, husat man	wanem, wanem sam- ting	long wanem taim	we	long wanem as	olsem wanem	hamas , hamas -pela	wanem	25% (2/8)

(Sources: David Sutcliffe and Peter Patrick, p.c., Sordam & Eersel 1985, Shanks 1994, Muysken & Smith 1990, Hancock 1987 and p.c., Verhaar 1995)

Table 13: *Percentage of wh-words shared among the English-based pidgins and creoles*

	Sranan	Djuka	Saramaccan	Krio	Cameroon	Tok Pisin
Jamaican	13% (1/8)	13% (1/8)	0% (0/7)	50% (4/8)	25% (2/8)	50% (2/8)
Sranan		100% (8/8)	29% (2/7)	0% (0/8)	13% (1/8)	0% (0/8)
Djuka			29% (2/7)	0% (0/8)	13% (1/8)	0% (0/8)
Saramaccan				13% (1/8)	25% (2/8)	0% (0/8)
Krio					50% (4/8)	25% (2/8)
Cameroon						0% (0/8)

in the earliest extant texts: for Sranan and Saramaccan we have included in parentheses the wh-words found in Herlein’s (1718) text and Riemer’s 18th century Saramaccan dictionary (cf. Arends and Perl 1995). Table 13 shows the percentages of wh-words that the creoles share with one another.

McWhorter (1995:326) posits the existence of a “central and robust” expanded pidgin English which influenced the development of all Atlantic creoles, and believes that this expanded pidgin English gave rise to Sranan, was transported to Jamaica, and then to Sierra Leone in the 1800s, which would have developed into modern Sierra Leone Krio (1995:323). Holm (1989), on the other hand, considers the typological similarity between the substrate Niger-Congo languages responsible for the similarities between the Atlantic English creoles. Clearly, though, the Atlantic creoles as a group have more wh-words in common among themselves than they have in common with a Pacific creole such as Tok Pisin, which had completely different substrate.

We note that Krio and Cameroonians seem to form a natural group, as do Sranan and Djuka. Jamaican retains all of the English wh-words, possibly due to input from settlers from Barbados.¹⁰

5.2. *Wh-constructions in English-based creoles*

In this section, we will advance predictions according to mutual linguistic accommodation, and then check these against the actual constructions.

Jamaican

By 1658, the white population in Jamaica was 4500 with only 1400 blacks, and seven years later, in 1664, about 1000 more settlers arrived in Jamaica from Barbados. Although there is no explicit record of these settlers bringing any slaves with them, it is likely that they did so. Between 1671 and 1675, around 1700 settlers came to Western Jamaica from Surinam, and according to Bilby

(1983:60) at least 980 of these were slaves. The fact that the wh-words in Jamaican are completely different from those in Surinam suggests that the Jamaican wh-words were stabilized before the arrival of new settlers. The large number of whites in the initial period of settlement and the fact that the slaves initially imported were from Barbados would account for the number of wh-words taken directly from English. The fact that both the African languages involved in the contact situation and English exhibit wh-word fronting would lead to the prediction, by mutual linguistic accommodation, that the wh-elements in Jamaican creole would also be fronted; this is attested: *WE(paat) im de* 'Where is he?' (Hancock 1987:285). One would not predict the presence of a focus particle based on mutual linguistic accommodation between standard colloquial English and the relevant African languages, because this variety of English did/does not possess anything comparable to the focus particle as do French and Portuguese. However, if we take into account that Hiberno-English may have had an influence in the formation of Barbadian, and thus also Jamaican Creole English, then there would be more reason to expect a focus particle. Interestingly, Jamaican does display the highlighter *a* that optionally appears before wh-words (*A wa Anti sen fi mi?* 'What has Auntie sent for me' [Bailey 1966:87]). Hancock (p.c.) points out that Yoruba as well as Hiberno-English display a type of construction similar to that found in Jamaican, which could be the source of *a* in questions by mutual linguistic accommodation:

Yoruba:

- (24) a. *Mo n lo.*
 1SG PROG go
 'I'm going.'
 b. *Li-l ni mo n lo.*
 COP-go FOC 1SG PROG go
 'I'm GOING!'

Hiberno-English:

- (25) a. *She's pretty.*
 b. *It's pretty she is.*

Sranan and Djuka

The first colony in Surinam was established by the British in 1651. The first hundred men sent to Surinam were mostly indentured servants from Barbados who wanted to start plantations of their own. In 1667, the Dutch took over, but

maintained the English-based creole. In the following three years, 1200 new slaves were imported in addition to the 1300 old slaves still there (Holm 1989:434).

The three major African languages that influenced Surinam were Kikongo, Gbe (Ewe, Fon) and Twi. Although the last two of these languages have a focus element in question formation, Sranan and Djuka do not. The reason for this may be that there is no comparable structure in colloquial English which could be used for this purpose, unlike in French and Portuguese, where the cleft construction was significantly more frequently found in question formation. However, if we assume again that Hiberno-English was involved in the formation of Sranan and Djuka, we would possibly expect a focus particle. In assessing a match between the construction in Hiberno-English and in Ewe (10) and Twi (13), we note that the Twi and Ewe focus particles appears after the *wh*-word, whereas in Hiberno-English it appears before the relevant element. By mutual linguistic accommodation, then, we would expect there to be less chance of the incorporation of a focus particle into Sranan question formation. This seems to be the case: we find no particle accompanying *wh*-words in Sranan.

Of course, given that the African languages in question had *wh*-word fronting, we predict fronting of *wh*-elements in Sranan, a correct prediction: *Ope a de* 'Where is he?' (Hancock 1987:285).

As mentioned above, Djuka, directly related to Sranan, is spoken by people whose ancestors escaped from the coastal plantations between 1715 and 1760. Given that Djuka was influenced by the same substrate languages as Sranan, it should also exhibit *wh*-word fronting and the absence of a focus particle. This prediction is borne out:

Djuka:

(26) a. *Pe a de?*

where he there

'Where is he?'

b. *Fa den sama e libi?*

how the people PROG live

'How do people manage to live?' (Hancock 1987: 285, 310)

Saramaccan

The groups speaking Saramaccan are descendants of slaves who escaped from Surinamese plantations in the 17th and 18th centuries. When Jewish and Dutch

plantation owners were forced out of Brazil by the Portuguese in 1664, they took with them slaves who probably spoke a Portuguese-based pidgin or creole. On their Surinam plantations developed a mixed English-Portuguese creole called “Djutongo” by contemporaries which was a precursor to Saramaccan (historical information from Holm 1989, Bakker, Smith & Veenstra 1995).

Given that all African languages involved in the formation of Saramaccan (most importantly Gbe (Ewe, Fon) and Twi exhibited wh-word fronting, while Hiberno-English had fronting and used a type of highlighting construction not structurally similar to the focus construction in Gbe and Twi, based on mutual linguistic accommodation one would expect wh-word fronting and most probably no focus particle in Saramaccan. On the other hand, if the African language influence was great enough, there may be more reason to expect a focus particle. However, Holm’s (1998:180) hypothesis that *na* is a possible candidate for focus particle is dismissed by Veenstra & Den Besten (1995:310), given that *na* is not an equative copula but an allomorph of a locative preposition; moreover, the particle *un*, another possible candidate, seems to be the equivalent of ‘which’.

Krio and Cameroonian

It was not until the late 1700s, when slavery was abolished in Britain, that Krio started to take its modern shape. In 1787, 330 blacks and 70 white prostitutes were sent to Sierra Leone from England. This was the beginning of Freetown. In 1792, 1100 former American slaves were brought to Freetown, and in 1800, 550 Jamaican Maroons. As a result of these events, Krio shares a number of features with other Atlantic creoles.

Subsequently, the British navy freed thousands of slaves from slave ships operated by other countries and released them in Freetown. Many settled in the town and learned Krio. Holm (1989) accounts for Krio structure by appealing to Niger-Congo languages; influence from Hiberno-English also cannot be ruled out. Interestingly, Krio possesses what Hancock calls a highlighter, an element consisting of copula *na* preceding the fronted wh-word (Hancock 1987:285, 311 and p.c.):

- (27) a. *Na usai yu dey go?*
 FOC where you PROG go
 ‘Where are you going?’

- b. *Na weytin duyon ebul du am?*
 FOC why you.NEG able do it
 'Why can't you do it?'

Note that the position of the particle *na* corresponds with that of the highlighter in Hiberno-English (25). However, use of the focus elements is optional (*Au pipil den de maneĵ to lib?* 'How do people manage to live' [Hancock 1987:310]).

The Krio highlighter can be easily accounted for by mutual linguistic accommodation: both Yoruba, one of the key languages in the formation of Krio, as well as Hiberno-English exhibit a similar construction. Apart from *wh*-word fronting, which is present and expected in Krio, we find *in situ* question formation as well, but it seems to have the same function as *in situ* question formation in English (Hancock, p.c.): fronting is the default option.

According to Holm (1989:430), it was the 18th century pidgin English used around Calabar on the Nigerian coast that later developed into Cameroonian. This pidgin was influenced by missionaries who started arriving after 1827. According to Todd (1984:94), 22 of these spoke Krio and at least six spoke Jamaican English. In 1884, Cameroon was taken over by the Germans, who treated Cameroonian as a distinct linguistic system and spread the language further inland.

The languages spoken by the communities which used the western variety of Cameroon Pidgin English have a focus particle and fronting, which would lead us to posit that *wh*-words would occur clause-initially. Since English also has fronting, mutual linguistic accommodation would lead us to predict fronting for Cameroonian as well, which is the case: *Ha pipo fit lif?* 'How do people manage to live?' (Hancock 1987:285). Cameroonian also allows *in situ* question formation, which, as in Krio, has a function similar to *in situ* question formation in English, and is thus not the default type. As for a highlighter or focus particle, we would expect Cameroonian to have one, since both the relevant African languages and Hiberno-English had one. However, we find no focus particle or highlighter in the Cameroonian data available to us. It may well be that because of the missionary influence in the early 19th century, the focus particle was lost, but this is simply conjecture.

Tok Pisin

Tok Pisin is comparable to Sranan in that after initial contact with English, it developed largely independently of its lexifier language. According to

Mühlhäusler (1982), early European settlers of the South Pacific used a Pacific pidgin influenced by Chinese Pidgin English to communicate with indigenes. In 1884 what is today Papua New Guinea became a protectorate of Great Britain, with Germany controlling the northeastern side of the island. However, pidgin English had already been established as a lingua franca in this section as well.

Tok Pisin was decisively influenced by the Melanesian-Austrolasan (MNAN) languages, which are VSO and generally use both *in situ* and fronting for wh-words. Among these languages, Tolai, also VSO, had the greatest influence on Tok Pisin. According to Goulden (1990), when most MNAN languages “agree on a way in which a syntactic category or semantic feature is realized at the lexical level, Tok Pisin often has a calque of the structure”. MNAN as VSO languages have both *in situ* and fronted wh-words. Although English has both, its default is wh-word fronting. Thus, one might expect Tok Pisin to have primarily wh-word fronting, and secondarily *in situ* question formation. From the data available to us, we were able to establish that Tok Pisin allows both options, but we were not able to ascertain whether fronting and *in situ* formation are largely interchangeable, or whether one is primary. The relevant examples, from Verhaar (1995:62), are given in (28).

- (28) a. *We stap papa bilong yu?*
 where be father your
 ‘Where is your father?’
 b. *Tok tru em i wanem samting?*
 talk truth he PM what
 ‘What is truth?’
 (PM = predicate marker)

MNAN languages do not have focus particles, and thus we would not predict Tok Pisin to have developed them under the hypothesis of mutual linguistic accommodation. Indeed, this is the case, as is apparent above.

Table 14 offers a summary of the findings from this section.

Predictions regarding the focus particle were correct six out of seven times (86%), while for fronting the predictions were found to be correct in all cases, which is not surprising since all languages involved display fronting as the clear default structure. Finally, frequent use of *in situ* question formation is found in Tok Pisin, which is expected given its MNAN primary substrate, Tolai.

Table 14: Predictions from mutual linguistic accomodation principles and attestations in English-based pidgins and creoles (predictions [P] and attested findings [Att.] are indicated in terms of presence [+] and absence [-]):

	Jamaican	Sranan	Djuka	Saramac.	Krio	Camer.	Tok P.
Focus	P: + Att: +	P: - Att: -	P: - Att: -	P: - Att: -	P: + Att: +	P: + Att: -	P: - Att: -
Fronting	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +	P: + Att: +
In situ	P: - Att: -	P: - Att: -	P: - Att: -	P: - Att: -	P: - Att: -	P: - Att: -	P: + Att: +

6. Conclusion and Final Remarks

In this study, one issue we have examined is the extent to which wh-words can be taken as a reliable diagnostic of relatedness between two pidgins or creoles. We have found that while common wh-words between two or more pidgins or creoles may reflect common development (Sranan and Djuka, Haitian and St. Lucian, Mauritian and Seychellois, or even Daman and Korlai Portuguese), it can also simply indicate that, as with Papiamentu and Palenquero, two creoles (or pidgins) simply share a common source language. We found that substrate influence can account unambiguously for the bimorphemic nature of the wh-words (Kristang), whereas in other cases (the Atlantic creoles and Daman and Korlai Portuguese), the bimorphemic nature of the wh-words was better accounted for by appealing to semantic transparency.

We have mainly examined, however, the assumption that the nature of question formation in pidgins and creoles can be accounted for by appealing to the interaction between the notions of mutual linguistic accommodation and universal tendencies. Our assumption was that if the languages in contact do not share a certain structure, then the structure of the question will be accounted for by universal tendencies such as the drive for semantic transparency.

This model achieved the correct predictions in just over 90% (92%, or 69/75) of the cases. Thomason & Kaufman (1988) have fused two views that have often been considered incompatible: substrate influence (either via compromise on a common structure or simple L1 transfer) and the bioprogram (universal tendencies). The results of the systematic study of one feature in 26 pidgin/creole languages strongly suggest that these two views are in fact

complementary, and that at this point in the development of pidgin and creole linguistics, to argue otherwise would be tantamount to disregarding the essential nature of language shift.

Notes

- * We are grateful to the following people who gave generously of their time in helping with the data compilation: Jacques Arends, Ian Hancock, Philippe Maurer, Peter Patrick, Ian Smith, and David Sutcliffe. Many thanks to John McWhorter and to an anonymous reviewer for their detailed editorial comments. Ahmar Mahboob contributed most of the section on the English-based creoles. Special thanks to Andrew Garboden for editorial assistance. J. Clancy Clements is responsible for remainder of the paper. All errors and other shortcomings are, of course, our own.

1. Baker (1995) presents a set of arguments in favor of the gradualist hypothesis, using data from a wide range of creoles. He argues that bimorphemic wh-words first developed in Mauritian Creole French (Mau) 101 years after its formation and in the Antillean French creoles 167 years after their formation.

Baker dates Mauritian to 1721 and notes that the first attestation of a bimorphemic interrogative was in 1822, 101 years later, while recognizing himself that there is only a handful of fragmentary texts from the 18th century. In Chaudenson (1981), we find various quotes of relevant 18th century French creole texts used by Baker. In all of them there is only one instance of a wh-word, *quand* used as a temporal conjunction in a 1799 text (1981:5). Thus the severe lack of 18th century texts may have something to do with Baker's dating. Another mitigating factor is that bimorphemic wh-words often existed together with monomorphemic wh-words. The texts available to us from Mauritian French are recorded by educated native speakers of French. Assuming that the consultants of the chroniclers were creole speakers, we do not know whether they were native speakers, nor do we know the circumstances under which the texts were collected. Finally, although it does not change the date significantly, we were able to attest a possible bimorphemic wh-word in a text Baker dates to 1818, where we find *qui dou-monde* 'who'. It is not unreasonable to believe that were more texts available, more bimorphemic wh-words would be attestable. Thus, the gradualist hypothesis is not strongly supported by this evidence.

Baker sets the date of formation for the Antillean French creoles at around 1635 and attests the first bimorphemic wh-word in 1802. We were able to consult only one of the two sources Baker uses for the 17th century texts, namely Rochefort (1658). In this volume, there is a Carib-French glossary that reveals nothing about bimorphemic wh-words. Again, the evidence for the gradualist hypothesis seems weak. In fairness, there is a small number of bimorphemic wh-words found in the 1802 texts. However, the texts are written dialogues between a white person and a black person (the former either a captain or a plantation owner, the latter either a deck hand or a foreman). Because of this, the same questions apply here as above with Mauritian.

2. There is disagreement in the literature about when Annobon was discovered. Bartens (1995:113) states it was discovered in 1417, which would be surely too early given that

Henry the Navigator, the famous Portuguese monarch who began the Portuguese colonial expansion, started his travels to North Africa around 1417. Cape Verde islands were discovered a little over 40 years later in 1462, and the other Gulf of Guinea islands (São Tomé and Príncipe) were only discovered around 1470-72. Barrena (1957:9) seems to have erred in the other extreme, putting Annobon's discovery in 1741. If, however, both 1417 and 1741 are considered typographical errors, which is what we believe, then both authors had the date 1471 in mind: "The Portuguese came upon the uninhabited island [i.e. Annobon] around 1471" (Holm 1989:283). The year 1471 is also plausible within the dates at which the other Gulf of Guinea islands were found (cf. also Post 1995:191).

3. It is not clear whether Kabuverdianu *pamo-di k'* and Kriyol *pabya dike ku* are related. If so, then these creoles would share 6/8 of their wh-words.
4. Sãotomense *keŋé ku* (< *ke* + *nge* [*nge* 'person'] + *ku* [< Ptg. *que*]) we take not to be related to KV *ken* (*ki*) or Kr *kin ku*, which derive from Portuguese *quem que* 'who'.
5. If Kriyol *kuma ku* is related to Angolar *m'ma*, then they share 3/8 of their wh-words.
6. We consider Principense *nĩgé ki ... â*, Ang *ngê*, Sãotomense *keŋé ku* and Annobonese *kenge* to be related in that they share the element *nge* 'person'. Moreover, Annobonese *xa /xa/* we take to be related to Principense *kwá ... â*, Sãotomense *kwa ku*, and Angolar *kwa* via a rule of velar stop fricativization. The same we claim is true of Annobonese *xafe*, which is related to Principense *kwe mǎda ki*, Sãotomense *kwa mǎda* and Angolar *kwa ma ... ra*. Of course, *xántu* is from Ptg. *quanto*, and finally, Annobonese *xamá* 'how' we take to be related to Angolar *m'ma*.
7. For quantitative comparisons among the French creoles, the forms *ki-sa*, *sa-ki*, etc. were counted all as one. In those cases in which we find new form and a form taken directly from French, both were counted.
8. The historical account given here is an amalgam of Kouwenberg and Muysken (1995), Bartens (1995), and Holm (1989).
9. The information on Palenquero was taken from Bartens (1995) and Holm (1989).
10. Holm (1989:470) emphasizes the influence of what he calls proto-Sranan, stating that at least 10% of the slaves at that linguistically critical time were from Surinam. The inventory of wh-words does not bear this influence out, though it does, as said, point to an influence from Barbados.

References

- Adu-Amankwah, D. and Botne, R. 1996. *A study guide to functional Twi*. Bloomington, IN: African Languages Program.
- Arends, J. and Perl, M. 1995. *Early Suriname creole texts: A collection of 18th century Sranan and Saramaccan documents*. Frankfurt: Vervuert.
- Arends, J., Muysken, P. and Smith, N. 1995. *Pidgins and creoles: An introduction*. Amsterdam: John Benjamins.

- Bailey, B. 1966. *Jamaican creole syntax*. Cambridge: Cambridge University Press.
- Baker, P. 1972. *Kreol: A description of Mauritian Creole*. London: C. Hurst and Company.
- Baker, P. 1995. "Some developmental inferences from historical studies of pidgins and creoles". In *The early stages of creolization*, J. Arends (ed.), 1-24. Amsterdam: John Benjamins.
- Bakker, P., Smith, N. and Veenstra, T. 1995. "Saramaccan". In Arends, Muysken and Smith (eds), 165-178.
- Barrena, N. 1957. *Gramática annobonesa*. Madrid: Consejo Superior de Investigaciones Científicas.
- Bartens, A. 1995. *Die iberoromanisch-basierten Kreolsprachen*. Frankfurt am Main: Peter Lang.
- Baxter, A. 1988. *A grammar of Kristang (Malacca Creole Portuguese)*. Canberra: The Australian National University.
- Bilby, K. 1983 "How the 'older heads' talk: A Jamaican Maroon spirit possession language and its relationship to the creoles of Suriname and Sierra Leone". *New West Indian Guide* 57: 37-88.
- Boakye, P. 1982. *Syntaxe de l'achanti: Du phoneme à la phrase segmentée*. Bern: Peter Lang.
- Bollée, A. 1977a. *Zur Entstehung der französischen Kreolendialekte im Indischen Ozean: Creolisierung ohne Pidginisierung*. Geneva: Librairie Droz.
- Bollée, A. 1977b. *Le créole français des Seychelles*. Tübingen: Max Niemeyer.
- Boretzky, N. 1983. *Kreolsprachen, Substrate und Sprachwandel*. Wiesbaden: Harrassowitz.
- Cardoso, E.A. 1989. *O crioulo da ilha de São Nicolau de Cabo Verde*. Lisboa: Casa da Moeda.
- Carrington, L.D. 1984. *St. Lucian Creole. A descriptive analysis of its phonology and morphosyntax*. Hamburg: Helmut Buske.
- Calvalhão Buescu, M.L. 1984. *Copilação de toda las obras de Gil Vicente*. 2 vols. Lisbon: Imprensa Nacional-Casa da Moeda.
- Cérol, M-J. 1991. *Un introduction au créole guadaloupéen*. Condé-sur-Noireau (France): Éditions Jasor.
- Chaudenson, R. 1981. "L'Atlas linguistique et ethnographique de Rodrigues: présentation et extraits". *Études Créoles* 4:117-50.
- Clements, J. C. 1996. *The genesis of a language*. Amsterdam: John Benjamins.
- Clements, J. C. 1992. "Elements of resistance in language contact-induced change". In *Explanation in Historical Linguistics*, G. Davis and G. Iverson (eds), 41-58. Amsterdam: John Benjamins.
- Corne, C. 1977. *Seychelles Creole grammar: Elements for Indian Ocean proto-creole construction*. Tübingen: TBL Verlag Gunter Narr.
- Da Silva Maia, A. 1964. *Licões de gramática de quimbundo*. Cucujães: Esc. Tip. Das Missões.
- Dalgado, S. R. 1917. "Dialecto indo-português de Negapatão". *Revista Lusitana* 20:40-53.
- Ehrhart, S. 1993. *Le créole français de St-Louis (le tayo) en Nouvelle-Calédonie*. Hamburg: Helmut Buske.
- Ferraz, L.I. 1974. "A linguistic appraisal of Angolar". In *In Memoriam António Jorge Dias*,

- 2, 177-86. Lisboa: Instituto de Alta Cultura/Junta de Investigações do Ultramar.
- Ferraz L.I. 1983. "The origin and development of four creoles in the Gulf of Guinea". In *The social context of creolization*, E. Woolford and W. Washabaugh (eds.), 120-5. Ann Arbor: Karoma.
- Ferraz, L.I. and Valkhoff, M.F. 1975. "A comparative study of São-tomense and Caboverdiano creole". In *Miscelânea Luso-africana: Colectânea de estudos coligidos*, M.F. Valkhoff (ed.), 15-39. Lisbon: Junta de Investigações Científicas do Ultramar.
- Gamble, D. P. 1991. *Elementary Gambian Wolof grammar*. Brisbane, CA: Gambian Studies No. 25.
- Givón, T. 1990. *Syntax: A functional-typological introduction* (Vol II). Amsterdam: John Benjamins.
- Gospel Missionary Union. 1967. *Bambara grammar*. (2 vols.) Bamako: Protestant Mission.
- Goulden, R. J. 1990. *The Melanesian content in Tok Pisin*. Canberra: The Australian National University.
- Günther, W. 1973. *Das portugiesische Kreolisch der Ilha do Príncipe*. Marburg an der Lahn: Selbstverlag.
- Hancock, Ian F. 1975. "Malacca Creole Portuguese: Asian, African or European?" *Anthropological Linguistics* 17:211-36.
- Hancock, Ian F. 1987. "A preliminary classification of the Anglophone Atlantic creoles, with syntactic data from thirty-three representative dialects". In *Pidgin and Creole Languages: Essays in memory of John E. Reinecke*, G.G. Gilbert (ed.), 264-334. Honolulu: University of Hawaii Press.
- Holm, J. 1988. *Pidgins and Creoles* (Vol. I). Cambridge: Cambridge University Press.
- Holm, J. 1989. *Pidgins and Creoles* (Vol. II). Cambridge: Cambridge University Press.
- Huber, J. 1933. *Gramática do Português antigo*. Lisbon: Fundação de Calouste Gulbenkian.
- Ittman, J. 1939. *Grammatik des Duala (Kamerun)*. Hamburg: Dietrich Reimer.
- Kihm, A. 1994. *Kriyol syntax*. Amsterdam: John Benjamins.
- Kouwenberg, S. and Muysken, P. 1995. "Papiamento". In Arends, Muysken and Smith, 205-18.
- Long, R. 1970. *Basic Bambara*. Bloomington: Indiana University Press.
- Maurer, P. 1995. *L'angolar*. Hamburg: Helmut Buske.
- McWhorter J. 1995. "Sisters under the skin: a case for genetic relationship between the Atlantic English-based creoles". *Journal of Pidgin and Creole Languages* 10: 289-333.
- Mühlhäusler, P. 1982. "Etymology and pidgin and creole languages". *Transactions of the Philological Society* 99-118.
- Muysken, P. and Smith, N. 1990. "Question words in pidgin and creole languages". *Linguistics* 28: 883-903.
- Neumann, I. 1985. *Le créole de Breaux Bridge, Louisiane*. Hamburg: Helmut Buske.
- Oger, L. 1982. *The English companion to Ibibemba Cakwa Chiti Mukulu*. Chinsali, Zambia: The Language Centre.
- Pasch, H. 1995. *Kurzgrammatik des Ewe*. Cologne: Rüdiger Köppe.
- Post, M. 1995. "Fa d'Ambu". In Arends, Muysken and Smith, 191-204.
- Rochefort, C. de. 1658. *Histoire naturelle et morale des Îles Antilles de l'Amérique*. Paris:

- Amout Leers.
- Seuren, P. and Wekker, H. 1986. "Semantic transparency as a factor in creole genesis". In *Substrata versus universals in creole genesis*, P. Muysken and N. Smith (eds), 57-70. Amsterdam: John Benjamins.
- Smith, N. 1995. "An annotated list of pidgins, creoles, and mixed languages". In Arends, Muysken and Smith, 331-74.
- Sordam, M. and Eersel, H. 1985. *Sranantongo/Surinaamse Taal*. Baarn, the Netherlands: Bosch & Keuning.
- Thomason, S. G. and Kaufman, T. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Todd, L. 1984. *Modern Englishes: Pidgins and creoles*. Oxford: Blackwell.
- Valdman, A. 1978. *Le créole: Structure, statut et origine*. Paris: Klincksieck.
- Veenstra, T. & Den Besten, H. 1995. "Fronting". In Arends, Muysken and Smith, 303- 15.
- Verhaar, W.M. 1995. *Toward a reference grammar of Tok Pisin: An experiment in corpus linguistics*. Honolulu: University of Hawaii Press.

Index

A

- acrolect 19, 31, 33, 66, 150, 178, 182
(See Creole continuum)
- Adjectives 58, 61, 63, 100, 101, 105,
110-113, 115-119, 160, 162, 167,
176, 178, 225
(See Attributive, Dynamic, Property
Items, Stative, Stativity)
- Adverbs 1, 2, 6-9, 12, 36, 61, 106, 107,
112, 113, 116, 162, 167, 189, 250
- Anglicization 31, 34, 36
- Angolar Creole Portuguese 65
- Anteriority 77, 78, 235
(See TMA markers)
- Aspect 24, 25, 35, 36, 38, 42, 53, 57,
62, 64, 66-68, 71-73, 75-78, 80, 81,
86, 91, 93, 98, 99, 106, 108, 109, 116,
117, 119, 131, 140, 145, 152, 157,
159, 160, 163, 179, 180, 182, 183,
188, 193, 194, 195-202, 212, 219-
221, 230-232, 234-236, 240, 242
(See Completive, Nonpunctual,
Perfect, TMA markers)
- Attributive 105, 112-116, 118
(See Adjectives, Property Items)

B

- Barlavento 16-23, 25
(See Cape Verdean Creole Portu-
guese, Sotavento)
- Basilect 19, 31, 33, 66, 149, 150, 178,
179
(See Creole continuum)
- Bemba 251
- Borrowing 41, 58, 63, 101, 116, 167,
193
- Brazilian Portuguese 22, 119, 120,
123-129

(See European Portuguese, Middle
Portuguese, Portuguese)

Bulgarian 58

C

- Cape Verde 5, 15-20, 27, 28, 30, 245,
258
- Cape Verdean Creole Portuguese
15-17, 19-21, 26, 28, 30, 65, 67, 69,
71, 73, 76, 78-81
(See Kabuverdianu cluster)
- Cayennais 150
- Chaos Theory 236-242
- Complementizer 37, 90, 91, 146, 153,
161
- Completive 25, 67, 69, 71, 75-78, 80,
91, 108, 109, 116, 145, 146, 196, 199,
200, 221, 228-232, 234
(See TMA markers)
- Componential diffusion 15, 17, 26
- Concord 119, 120, 122-125, 127, 130
(See Inflection, Morphology)
- copula 4-6, 14, 25, 26, 61, 75, 80, 81,
106, 108, 110-112, 114, 116-118,
160, 163, 165, 178, 180, 256
- Creole continuum 31, 33, 48, 49, 81,
118, 194, 235, 240-242
(See Basilect, Decreolization,
Mesolect, Acrolect)
- Creole English 46, 49, 65, 66, 101, 130,
151, 255
- Creole genesis 19, 30, 81, 82, 85, 95,
100, 117-119, 132, 140, 147, 152,
153, 171, 178-180, 193, 194, 196,
236, 237-240, 242, 243, 260
- Creolization 15-20, 22, 26, 29, 46,
48-52, 57, 60-62, 65, 82, 85, 86, 88,

92, 95, 98-100, 118-120, 126, 127, 151, 152, 153, 164, 169, 171, 178, 180, 193, 234, 238, 242-244, 247, 259, 260
 Decreolization 19, 20, 23, 28, 46, 57, 80, 159, 179, 243, 248
 (See Creole continuum)
 Diagnostics 1, 4
 Diglossia 19, 20, 27, 29, 31, 32, 48
 Dominican Creole French 65
 DP 87, 91, 248-250
 Drift 52, 64, 119, 128
 dynamic 202, 220, 221, 226, 237, 238
 (See Adjectives, Stative, Stativity)

E

Early Romance 49-51, 53, 56, 60, 61, 64
 English 2, 3, 5, 9, 10, 13, 15, 21, 28, 30-49, 52, 56-58, 64-66, 72, 75, 77, 80, 82, 84, 85, 88, 89, 91, 92, 96, 97, 101, 102, 104, 105, 116-119, 130, 136-145, 147, 149-155, 158, 160, 167, 168, 171, 172-180, 182-184, 186, 187, 190-195, 197, 198, 204, 209, 213, 216, 219, 221, 222, 226, 228, 229, 232, 234-236, 243, 244, 251, 253-258, 260
 (See Middle English, Old English)
 European Portuguese 22, 119-122, 126
 (See Brazilian Portuguese, Middle Portuguese, Portuguese)
 Event hypothesis 132, 133, 136, 138, 149, 150

F

Floating quantifiers 1, 6-8, 10, 11, 15
 Focus 1, 33, 74, 75, 80, 88, 106, 117, 147, 182, 205, 208, 209, 211, 232, 233, 243, 247, 248, 251, 252, 254, 255-257
 French 1, 2, 5, 7, 9, 15, 28, 31, 33, 34, 36-38, 40-43, 45-51, 53-57, 60-65, 80, 84, 87, 89-92, 94, 96-99, 102, 103, 104, 132, 146, 150-158, 160-

162, 164, 165, 167, 168, 177, 184, 187, 190, 193, 221, 232, 237, 242-244, 250-256, 258, 259

Future 5, 19, 20, 23, 30, 35, 60, 61, 63, 66, 80, 91, 127, 140, 145, 159, 194, 197, 200, 201, 203, 204, 212, 213, 214-220, 224, 228, 232-234, 237
 (See TMA markers)

G

Gallo-Romance 57
 (See Provençal)
 Gbe languages 115
 Germanic 15, 52, 53, 56, 57, 60, 64, 100
 Gradualist hypothesis 258
 Grammaticalization (Grammaticization) 60, 61, 81, 161, 179-183, 191-194, 220, 229, 232, 235
 Greek 49, 52-65, 85, 88, 96, 137
 Guyanese Creole English 65

H

Habitual 24, 25, 68, 71-73, 78, 80, 145, 160, 197, 198, 200, 204, 220, 221, 223, 225
 (See TMA markers)
 Haitian Creole French 65
 Hawaiian language 142, 143
 Hiberno-English 102, 255-257
 "High" Kwéyòl 31, 38
 (See Kwéyòl)
 Historical documentation 122, 128, 165
 History (of colonies) 3, 15, 16, 32, 34, 36, 38, 42, 48, 51-53, 60-64, 85, 98, 102, 120, 126-128, 153-157, 165, 167, 168, 170, 175, 176, 178, 179, 184, 192, 194, 235-239, 242
 History of language 36

I

Icelandic 2, 7, 9, 10, 15, 57, 65
 Imperfective 80, 108, 159, 196, 198, 200, 201, 204, 213, 218-222, 225, 226, 228, 232, 234

(See TMA markers)
 Inflection 2, 5, 10, 12, 13, 15, 64, 72,
 80, 99, 164
 (See Concord, Morphology)
 Information structure 180, 188, 190
 Instrumentalization 43, 44, 47
 Intellectualization 43, 44, 48
 Interference through shift 101
 Irrealis 23-25, 80, 145, 159, 160, 186,
 195, 196, 213, 234, 236
 (See TMA markers)
 Isomorphism 83-87, 92-94, 96

J

Jamaican Creole English 65, 255
 Japanese immigrants 148

K

Kabuverdianu cluster 15, 18, 20
 (See Cape Verdean Creole Portuguese)
 Krio Creole English 65, 66, 68-70, 72,
 74, 77, 79, 80, 255-257
 Kru languages 15, 175, 177
 Kwéyòl 31, 33-47
 (See "High" Kwéyòl)

L

Language attitudes 144
 Language attrition 144
 Language Bioprogram Hypothesis 29,
 82, 85, 99, 152, 239
 (See Protolanguage)
 Language ideology 35, 36, 48, 49
 Language policy 17, 19, 20
 Langue véhiculaire 154, 155
 Latin 49-65, 120, 127, 170, 241, 242,
 247
 Lexicon 26, 28, 32-37, 42-48, 79,
 83-86, 88-90, 94-96, 98, 100,
 150-152, 158-163, 167, 171, 179,
 195, 232, 233, 243, 248
 Liberian Settler English 171, 172, 179
 Língua Geral 120

M

Mainland Scandinavian 3, 9, 10
 (See Swedish)
 Malay 248, 249
 Mande languages 16, 166, 173, 175,
 177
 Mauritian Creole French 258
 mesolect 19, 193, 194
 (See Creole continuum)
 MF 83-88, 92, 94, 97, 98
 Middle English 3, 52
 Middle Portuguese 247
 Mood 24, 25, 28, 36, 58, 59, 61, 66,
 80-82, 106, 116, 127, 159, 160, 179,
 195, 196, 199, 201, 213, 233, 234,
 235, 236, 242
 (See TMA markers)
 Morphology 1-3, 8, 13, 15, 50, 53,
 56-61, 63, 83, 84, 86, 92, 98-101,
 154, 157, 159, 163, 166, 176, 179,
 182, 183, 184, 186, 187, 189, 192
 (See Concord, Inflection)
 Morphophonological Form 83
 morphosyntax 15, 22, 30, 57, 62, 259
 Multivariate analysis 189, 190

N

Nagamese 65, 68, 70, 72, 74, 76, 79, 80
 Narrative structure 183, 184, 186, 188,
 189
 Nativization 120, 126-128, 130-134,
 136, 139, 140, 142, 145, 147, 148,
 151, 153, 164, 170, 194
 Ndyuka Creole English 88, 99-103,
 105-118
 Negation 1, 2, 4, 11, 25, 28, 49, 92,
 106, 160, 163, 167
 Negerhollands Creole Dutch 65
 Nigerian Pidgin English 179, 180,
 192-194
 Nonpunctual 36, 195
 (See TMA markers)
 Nubi Creole Arabic 65

O

- Old English 2, 56, 64, 179
- Old French 55-57
- Old Irish 56, 60, 64
- Optimality Theory 91, 96, 171, 173, 178, 179
- orthography 19, 20, 38, 39, 44, 47, 48

P

- Palenquero Creole Spanish 65
- Papiamentu Creole Spanish 65
- Passive 23, 39, 41, 59-61, 63, 80, 143, 163
- Perfect 23, 58, 60, 61, 74, 76, 77, 82, 83, 85, 86, 88, 92, 93, 95, 96, 156, 197, 200, 203, 230-232, 234
(See Aspect)
- phonetics 20, 179
- phonology 20, 47, 58, 152, 154, 157, 159, 170-173, 175, 179, 240, 259
(See Optimality Theory)
- Pidgin 17, 28-32, 47-49, 62, 64, 65, 80-82, 89-91, 93-95, 97-100, 102, 104, 105, 116-120, 126, 127, 130, 131, 132, 136-139, 143-172, 175, 177-183, 192-194, 228, 235, 236, 238, 241-244, 252, 253, 254-258, 260
- Pidginization 17, 29, 48, 52, 65, 98, 99, 120, 127, 152, 153, 157-159, 162, 164, 165, 168-170, 178, 244
- Portuguese 1, 11, 13-28, 30, 54, 55, 60, 65, 73, 79-81, 91, 98, 99, 104, 119-130, 134-139, 141-145, 147, 148, 149-151, 221, 228, 229, 242-251, 253-256, 258-260
(See Brazilian Portuguese, European Portuguese, Middle Portuguese)
- Predicate 10, 40, 80, 82, 85, 106, 110, 111, 116-119, 126, 159, 178, 198, 201, 220, 221, 232, 235, 243, 257
- Predicate clefting 40
- Property items 100, 101, 105-119, 225, 234, 236
(See Adjectives, Attributive))
- Protolanguage 85, 86, 89-95, 97, 99

(See Language Bioprogram Hypothesis)

- Provençal 54-57, 61
(See Gallo-Romance)

Q

- Quantifier float 7

R

- Reduplication 40, 112, 118, 160, 179
(See Adjectives, Attributives)
- Register 31-34, 37, 39-43, 45, 47, 48
(See Style)
- Relative past 196, 200, 203, 204, 212, 217, 232
(See TMA markers)
- Relative tense 181, 182, 191, 204, 217, 228
(See TMA markers)
- Rumanian 54, 57, 60, 63

S

- Sango 130, 153-171
- Seychellois Creole French 65
- Slaves 16, 17, 22, 27, 38, 50, 102-105, 117, 131, 132, 150, 151, 155, 156, 244, 245, 248, 251-256, 259
- Sotavento 16-26, 28
(See Barlavento, Cape Verdean Creole Portuguese)
- Spell Out 83, 84, 86, 91
- Sranan Creole English 65
- St. Lucia 31-38, 43-48, 252
- St. Lucian English 33, 46-48
- Standardization 20, 44, 45
- Stative 14, 66-80, 108, 109, 116, 118, 180, 182, 198, 201, 204, 221, 225, 226, 231
(See Adjectives, Attributive, Dynamic, Stativity)
- Stativity 65, 75, 79, 80, 117, 183, 221
(See Adjectives, Attributive, Dynamic, Stative)
- Strong 1, 2, 13, 27, 51, 56, 61, 80, 83, 87, 97, 105, 125, 131, 142, 150, 184,

187, 191, 199, 216, 232, 252
 Style 32, 40, 44, 48, 49, 62, 121, 183, 194
 (See Register)
 Subject/verb number 119
 Substratal influence 17, 28, 62, 99, 101, 102, 117, 145, 148, 150, 153, 171, 172, 173, 177, 195, 244, 248, 250, 253, 254, 258
 Suppletion 85, 88, 96-98
 Surinamese Plantation Creole 100, 101
 (See Sranan Creole English, Ndyuka Creole English)
 Swedish 2, 3, 104
 (See Mainland Scandinavian)
 Syntax 1, 9, 14, 15, 30, 47, 61, 65, 66, 79-81, 83-90, 92, 93, 96-101, 105, 112, 118, 119, 127, 145, 150, 151, 152-154, 156, 158-160, 181, 192-194, 232, 235, 236, 241, 242, 259, 260

T

Tamil 248, 250, 251
 Target language 244
 Tense 2, 3, 5, 7, 9, 13, 23, 36, 58-61, 64, 66-81, 84, 91, 98, 99, 106, 116, 118, 119, 121, 127, 131, 145, 149, 152, 159, 160, 163, 179-184, 186, 187, 191-201, 204, 212-214, 217, 218, 220, 228, 232, 234-236, 242

(See Future, Relative Tense, TMA markers)

TMA markers 4, 6, 9, 10, 68, 70, 71, 73, 108, 111, 145
 (See Anteriority, Aspect, Completive, Habitual, Irrealis, Mood, Nonpunctual, Relative Past, Relative Tense, Tense)
 Tok Pisin Pidgin 65
 Tok Pisin Pidgin/Creole English 65
 Tones 29, 159, 160, 166
 typological distance 244, 249

U

Ubangi river basin 156, 165
 Unmarked verbs 66, 67, 71, 79, 109, 183, 198, 201-203, 205, 206, 208, 210, 211, 233, 234

V

Vernacular English of St. Lucia 33
 (See VESL)
 Vernacular Liberian English 171, 179
 VESL 33, 35-37, 40, 42, 45, 46
 (See Vernacular English of St. Lucia)

W

weak 2, 13, 61, 258
 Welsh 58

In the CREOLE LANGUAGE LIBRARY series the following volumes have been published thus far or are scheduled for publication:

1. MUYSKEN, Pieter & Norval SMITH (eds): *Substrata versus Universals in Creole Genesis. Papers from the Amsterdam Creole Workshop, April 1985*. 1986.
2. SEBBA, Mark: *The Syntax of Serial Verbs: an Investigation into Serialisation in Sranan and other languages*. 1987.
3. BYRNE, Francis: *Grammatical Relations in a Radical Creole: Verb Complementation in Saramaccan. With a foreword by D. Bickerton*. 1987.
4. LIPSKI, JOHN M.: *The Speech of the Negros Congos in Panama*. 1989.
5. JACKSON, Kenneth David: *Sing Without Shame. Oral traditions in Indo-Portuguese Creole verse*. 1990.
6. SINGLER, John V. (ed.): *Pidgin and Creole Tense/Mood/Aspect Systems*. 1990.
7. FABIAN, Johannes (ed.): *History from Below. The 'Vocabulary' of Elisabethville by André Yav: Text, Translations and Interpretive essay*. 1990.
8. BAILEY, Guy, Natalie MAYNOR and Patricia CUKOR-AVILA (eds): *The Emergence of Black English: Text and commentary*. 1991.
9. BYRNE, Francis and Thom HUEBNER (eds): *Development and Structures of Creole Languages. Essays in honor of Derek Bickerton*. 1991.
10. WINFORD, Donald: *Predication in Carribean English Creoles*. 1993.
11. BYRNE, Francis & John HOLM (eds): *Atlantic Meets Pacific: A global view of pidginization and creolization. Selected papers from the society for Pidgin and Creole Linguistics*. 1992.
12. BYRNE, Francis & Donald WINFORD (eds): *Focus and Grammatical Relations in Creole Languages: Papers from the University of Chicago Conference of Focus and Grammatical Relations in Creole Languages*. 1993.
13. ARENDS, Jacques (ed.): *The Early Stages of Creolization*. 1995.
14. KIHM, Alain: *Kriyol Syntax. The Portuguese-based Creole language of Guinea-Bissau*. 1994.
15. ARENDS, Jacques, Pieter MUYSKEN and Norval SMITH (eds): *Pidgins and Creoles. An introduction*. 1995.
16. CLEMENTS, J. Clancy: *The Genesis of a Language: The formation and development of Korlai Portuguese*. 1996.
17. THOMASON, Sarah G. (ed.): *Contact Languages. A wider perspective*. 1997.
18. ESCURE, Genevieve: *Creole and Dialect Continua. Standard acquisition processes in Belize and China (PRC)*. 1997.
19. SPEARS, Arthur K. and Donald WINFORD (eds): *The Structure and Status of Pidgins and Creoles*. 1997.
20. RICKFORD, John R. and Suzanne ROMAINE (eds.): *Creole Genesis, Attitudes and Discourse. Studies celebrating Charlene J. Sato*. 1999.
21. McWHORTER, John (ed.): *Language Change and Language Contact in Pidgins and Creoles*. 2000.
22. NEUMANN-HOLZSCHUH, Ingrid and Edgar W. SCHNEIDER (eds.): *Degrees of Restructuring in Creole Languages*. 2000.